

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2005, 06:24:33 ; Search time 15452 Seconds
(without alignments)
11499.177 Million cell updates/sec

Title: US-10-600-070-3
Sequence: 1 tgcctcgaataaggagat.....ctatacataaggctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb ba:
2: gb hcg:
3: gb in:
4: gb om:
5: gb ov:
6: gb pat:
7: gb ph:
8: gb pl:
9: gb pr:
10: gb ro:
11: gb sts:
12: gb sy:
13: gb un:
14: gb vl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3664	99.9	3668	8	AY221468
2	3660.8	99.8	3664	8	AY221467
3	3659	99.8	85791	8	AB016888
4	1931	52.7	2637	8	AY091075
5	1752.8	47.8	2438	8	AY221469
6	1744.6	47.6	2436	8	AY150490
7	317.6	8.7	150462	8	AP004885
8	304.4	8.3	341	8	AY199896
9	57.6	1.6	7218	6	166494
10	54	1.5	1141	6	AK083744
11	54	1.5	250029	3	AE014820
12	53.8	1.5	205796	2	AL713891
13	53.8	1.5	215524	10	AL603913
14	53.8	1.5	250078	3	AE014829
15	53.8	1.5	333321	3	AC116986
16	53.4	1.5	53391	8	AY199897
17	53.2	1.5	198573	2	AC110118
18	53	1.4	184039	10	AC139376
19	53	1.4	233373	2	AC134805

20	52.6	1.4	222859	2	AC119799	AC119799	Pan trogl
21	52	1.4	177010	2	AC140372	AC140372	Mus muscu
22	52	1.4	245924	5	BX004773	BX004773	zebrafish
23	52	1.4	255036	10	AC118206	AC118206	Mus muscu
24	52	1.4	268597	2	AC094787	AC094787	Rattus no
25	51.6	1.4	2000	6	AK655393	AK655393	Sequence
26	51.6	1.4	23213	9	AL590436	AL590436	Human DNA
27	51.6	1.4	110000	2	AC108609	AC108609	Continuation (3 of
28	51.6	1.4	292533	2	AC107260	AC107260	Rattus no
29	51.4	1.4	139995	10	AC11797	AC11797	Mus muscu
30	51.4	1.4	150579	10	AC126937	AC126937	Mus muscu
31	51.4	1.4	171307	9	AC104020	AC104020	Homo sapi
32	51.4	1.4	193634	2	AC068495	AC068495	Mus muscu
33	51.4	1.4	253094	2	AC113774	AC113774	Rattus no
34	51.2	1.4	227202	2	AC125636	AC125636	Rattus no
35	51.2	1.4	253772	2	AC109777	AC109777	Rattus no
36	51.2	1.4	313096	2	AC131879	AC131879	Rattus no
37	51.2	1.4	339033	2	AC106969	AC106969	Rattus no
38	51	1.4	202358	10	AL596187	AL596187	Mouse DNA
39	51	1.4	269375	2	AC114367	AC114367	Rattus no
40	50.8	1.4	110000	2	AC120236	AC120236	Rattus no
41	50.8	1.4	196519	2	AC127834	AC127834	Rattus no
42	50.8	1.4	266661	2	AC126523	AC126523	Rattus no
43	50.8	1.4	316828	2	AC135386	AC135386	Rattus no
44	50.6	1.4	106693	8	AP004188	AP004188	Oryza sat
45	50.6	1.4	144652	8	AC132003	AC132003	Oryza sat

ALIGNMENTS

RESULT 1
LOCUS AY221468 3668 bp DNA linear PLN 05-AUG-2003
DEFINITION Arabidopsis thaliana division protein (ARCE) gene, complete cds;
ACCESSION AY221468
VERSION AY221468.1 GI:33436338
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 3668)
OsterYoung,K.W.
OsterYoung,K.W.
TITLE ARCE is a J-Domain Plastid Division Protein and an Evolutionary
JOURNAL Descendant of the Cyanobacterial Cell Division Protein Ftn2
PUBMED 12897262
REFERENCE 2 (bases 1 to 3668)
Viltha,S., Koksharova,O., van Exp,H., Froehlich,J.E. and
OsterYoung,K.W.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) Department of Plant Biology, Michigan State
UNIVERSITY, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
source
1..3668
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/ecotype="Wassiljevskija"
478..3515
/gene="ARC6"
/note="synonym: At5g42480"
join(478..984,1067..1291,1379..1639,1723..2710,
3000..3133,3225..3515)
/gene="ARC6"
/product="division protein"
join(478..984,1067..1291,1379..1639,1723..2710,3000..3133,
3225..3515)
/gene="ARC6"

/note="inner envelope membrane-localized; similar to
cyanobacterial cell division protein sl10169 of
Synchocystis; contains J-domain"
/codon_start=1
/product="division protein"
/protein_id="AAQ18645.1"
/db_xref="GI:33436339"
/translation="MEALSHVIGIGLSPOLCRLPPTATKLRSHSTTTI CSASKMWAD
RLSDNFPTSSSSSPATATTAATLVLSPSIDRPERHATPIDFQVGAQTHFT
RLDRAPEARVSKPQGFDDALISRLLOAACEIDSPRSRREVEQLADEBT
VITVDPRDKYRGALCVTQEGREIYRVEBALIKERLPSRPFQDVLTALAFIDYS
RDMALDPRDPTGYEFEBEALKILQREGASSLAPDLRAQIDLETLEITPRYVLELLG
LPGLADYAAKRLNGLSGRNLIWSVGGGASALVGLTRKFMNEALIKRTAAQVQL
FVATPSNIPASEFEVEVALALVAQFI GKQPHILQDAPQFOQLQAKWAMEIPAM
LYDTRNWEIDFGLERGLCALLIGKVDCEMWLGLDSDSQYRPAIVERFLENSND
DNNDLPGLCKLLETWLAGVPPRPDTRDKKFKLDYDDPMTLSYERFVQGSPL
AAATARIKAEHYKASAMQALOKVPPSRITDRSNAPRKYQVTFVSDVGNVGR
GEFGEVFLAEVRSENFEINDVATIRGVSSSVDETIVENSADMLKASVKLIAGV
AIGLISLFSQKYLKSSSFQKQDWSMSDVATIGSVRADSEALPRMDARTAEI
VSKWOKIKSLAFGPDHRIEMLPVLDGRMLKITDRAAETRAELGLVDYTLKLSVDS
VTVSADOTRALVEATLEESACLSLVHPENNATDVRTYTRVEVFMKSGMKITGSEV
LAS"

ORIGIN

Query Match 99.9%; Score 3664; DB 8; Length 3668;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 TCTGCATTAGAGAAATACATTTTAGCAATTTGTCTTGATTTTCAACAAGATTTCCT 63
DB 1 TCTGCATTAGAGAAATACATTTTAGCAATTTGTCTTGATTTTCAACAAGATTTCCT 60
QY 64 GGCCTATGAGATTCATGGCTCTGTTTCTTTTCAATTTACATGTCATATAGTTTCAAT 123
DB 61 GGCCTATGAGATTCATGGCTCTGTTTCTTTTCAATTTACATGTCATATAGTTTCAAT 120
QY 124 TTACACATTTCAAGTTGATGTTTAAAGAAAGAGGAAATTGATGGGCTTTTGTGGTTT 183
DB 121 TTACACATTTCAAGTTGATGTTTAAAGAAAGAGGAAATTGATGGGCTTTTGTGGTTT 180
QY 184 AAATTTAAAGTGTCAAGAAATTAAGTCATTTGTTTCTGTTGCTCAATATGTTAAAT 243
DB 181 AAATTTAAAGTGTCAAGAAATTAAGTCATTTGTTTCTGTTGCTCAATATGTTAAAT 240
QY 244 GAAGGCAACTCCAAACGGTTCTTAGTGGAATAGATTATTAGACATTTAAACATCAATA 303
DB 241 GAAGGCAACTCCAAACGGTTCTTAGTGGAATAGATTATTAGACATTTAAACATCAATA 300
QY 304 GTCGCTGCGACGTGAATCATATGATGTTTTTATTTTTTTCAGTAGCTGTGATGTT 363
DB 301 GTCGCTGCGACGTGAATCATATGATGTTTTTATTTTTTTCAGTAGCTGTGATGTT 360
QY 364 TTTTGAATTTAATTATCTACTCAAAATCAAAATTTCAATTAACCTTGAGAGCAACAAACAG 423
DB 361 TTTTGAATTTAATTATCTACTCAAAATCAAAATTTCAATTAACCTTGAGAGCAACAAACAG 420
QY 424 TCTCTTAATATGTAAACAGAAACAAGTTTTTTTGTAGAGCTTAAAGACACTCCCATG 483
DB 421 TCTCTTAATATGTAAACAGAAACAAGTTTTTTTGTAGAGCTTAAAGACACTCCCATG 480
QY 484 GAAAGCTCTGAGTCAAGTCGGGCAATTTGTCCTCCCATTTCAATTAATGACCAATGCAACG 543
DB 481 GAAAGCTCTGAGTCAAGTCGGGCAATTTGTCCTCCCATTTCAATTAATGACCAATGCAACG 540
QY 544 GCGACGCAAAAGCTCCGACGTAGCCACAACACTCTACACTATCTGCTCGCCAGCAAA 603
DB 541 GCGACGCAAAAGCTCCGACGTAGCCACAACACTCTACACTATCTGCTCGCCAGCAAA 600
QY 604 TGGGCGCAACGGTCTTCTCTCGACCTTCAATTTCACTCCGATTTCTCTCTCTCTTTC 663
DB 601 TGGGCGCAACGGTCTTCTCTCGACCTTCAATTTCACTCCGATTTCTCTCTCTCTTTC 660
QY 664 GCCACCGCAACACACCGGCACTCTGCTCTGCGACACATCTATTTGATGTCGCCGAA 723

```

```

DB 661 GCCACCGCAACACACCGGCACTCTGCTCTGCGACACATCTATTTGATGTCGCCGAA 720
QY 724 CGCAGCTCCCAATCCCAATTTGATTTCTACCAAGTATTTAGAGCTCAAAACAATTTCTTA 783
DB 721 CGCAGCTCCCAATCCCAATTTGATTTCTACCAAGTATTTAGAGCTCAAAACAATTTCTTA 780
QY 784 ACCGATGGAATCAAGAAAGCAATTCGAAGCTAGGGTTTCGAAACCGCCGCAATTCGGTTTC 843
DB 781 ACCGATGGAATCAAGAAAGCAATTCGAAGCTAGGGTTTCGAAACCGCCGCAATTCGGTTTC 840
QY 844 AGGAGAGCGCTTTAATCAAGCCGAGACAGATTTCTCAAGCTCTGCGAAACCTGTCT 903
DB 841 AGGAGAGCGCTTTAATCAAGCCGAGACAGATTTCTCAAGCTCTGCGAAACCTGTCT 900
QY 904 AATCTCGGCTGAGAAAGAGTACAAATGAAGTCTTCTTGATATGAAGAGCTACAGTC 963
DB 901 AATCTCGGCTGAGAAAGAGTACAAATGAAGTCTTCTTGATATGAAGAGCTACAGTC 960
QY 964 ATCACTGATGTTCTTGGGATTAAGTTATTCGATTTCCGAATTAATTAAGTTTCTTCTGTT 1023
DB 961 ATCACTGATGTTCTTGGGATTAAGTTATTCGATTTCCGAATTAATTAAGTTTCTTCTGTT 1020
QY 1024 TTAATTTCAATGATTTGATTAAGAAAGAACTTTTATCTATGTAAGTTCCTGCGGCTCT 1083
DB 1021 TTAATTTCAATGATTTGATTAAGAAAGAACTTTTATCTATGTAAGTTCCTGCGGCTCT 1080
QY 1084 CTGCTGATTTGCAAGAAAGTGTGAGACTGAGATAGTTCTTCCGGTTGTGAGGCTCTGCT 1143
DB 1081 CTGCTGATTTGCAAGAAAGTGTGAGACTGAGATAGTTCTTCCGGTTGTGAGGCTCTGCT 1140
QY 1144 TAAAGAGAGGTTGCTTAAGTGTGTTTAAAGCAAGTGTGTTTAAAGTGTGCTGCTGCT 1203
DB 1141 TAAAGAGAGGTTGCTTAAGTGTGTTTAAAGCAAGTGTGTTTAAAGTGTGCTGCTGCT 1200
QY 1204 TCTCGATGTTCTGAGGAGTGTATGTCATTTGATTTCAACCTGATTTTATTAATGTTATGA 1263
DB 1201 TCTCGATGTTCTGAGGAGTGTATGTCATTTGATTTCAACCTGATTTTATTAATGTTATGA 1260
QY 1264 GTTTGTTGAGAGAGCTTTGAAAGCTTTTACAGTAGTTTGACTTGTGTTGTTGTTGTTGACG 1323
DB 1261 GTTTGTTGAGAGAGCTTTGAAAGCTTTTACAGTAGTTTGACTTGTGTTGTTGTTGTTGACG 1320
QY 1324 AGCGTTGCTTTTAAAGAACTTTCTTGATTTGATTAATCTTTGTTATTTGATCTTGTGTAAGA 1383
DB 1321 AGCGTTGCTTTTAAAGAACTTTCTTGATTTGATTAATCTTTGTTATTTGATCTTGTGTAAGA 1380
QY 1384 GGAAGAGCAAGTAAAGCTTGACACCGGATTTACGTGCAAAATTTGATGAGAGCTTTGAGAGA 1443
DB 1381 GGAAGAGCAAGTAAAGCTTGACACCGGATTTACGTGCAAAATTTGATGAGAGAGCTTTGAGAGA 1440
QY 1444 GATCACTCCGCTTATGTTCTTGAGAGCTAATTTGAGCTTACCGCTTGATGATTAACGCTGC 1503
DB 1441 GATCACTCCGCTTATGTTCTTGAGAGCTAATTTGAGCTTACCGCTTGATGATTAACGCTGC 1500
QY 1504 GAAAGACATTAATGTTTAAAGCGGTGTGCGGAATTTTGTGTTCTGTTGAGAGAGGTG 1563
DB 1501 GAAAGACATTAATGTTTAAAGCGGTGTGCGGAATTTTGTGTTCTGTTGAGAGAGGTG 1560
QY 1564 AGCATCAGCTCTTGTGGGGGTTTGAACCGTGAGAGTTTAATAAGAGCGGTTTTTACG 1623
DB 1561 AGCATCAGCTCTTGTGGGGGTTTGAACCGTGAGAGTTTAATAAGAGCGGTTTTTACG 1620
QY 1624 AATGACAGCTGCTGAGCAGGTATACAGTTTAAATCTTTTAAATTTCTTTAGCATGA 1683
DB 1621 AATGACAGCTGCTGAGCAGGTATACAGTTTAAATCTTTTAAATTTCTTTAGCATGA 1680
QY 1684 TATTAATTTAGTTTCTCATTTTAAATGATGTTGTGTGAGTGAATCTTTTGTAGCT 1743
DB 1681 TATTAATTTAGTTTCTCATTTTAAATGATGTTGTGTGAGTGAATCTTTTGTAGCT 1740
QY 1744 ACCCCAGCAATATTCAAGAGAGTCAATTTGAAGTTTACAAATTTGCACTTGTCTTGTG 1803

```

Db 1741 ACCCCAGACATATTCAGCAGAGTCATTTGAGATTACGAGTTCAGCTTCTGTG 1800
 Qy 1804 GCTCAAGCTTTATTTGGTAAGAGCCACCTTTTACAGATGCGATTAAGCAATTCAG 1863
 Db 1801 GCTCAAGCTTTATTTGGTAAGAGCCACCTTTTACAGATGCGATTAAGCAATTCAG 1860
 Qy 1864 CAATTCAGCAGGCTTAAGTAAATGCTATAGAGATTCCTGCAATGTTATATGATACAG 1923
 Db 1861 CAATTCAGCAGGCTTAAGTAAATGCTATAGAGATTCCTGCAATGTTATATGATACAG 1920
 Qy 1924 AATTAATGGAGATAGACTTGGTCTAGAAAGGGAGCTGTGCACTGCTTATAGGCAA 1983
 Db 1921 AATTAATGGAGATAGACTTGGTCTAGAAAGGGAGCTGTGCACTGCTTATAGGCAA 1980
 Qy 1984 GTTGAATAGTCGATAGTGTGGGCTTAGACATGAGATTCACATATAGGAATCCA 2043
 Db 1981 GTTGAATAGTCGATAGTGTGGGCTTAGACATGAGATTCACATATAGGAATCCA 2040
 Qy 2044 GCTAATGGAGTGTGTTTGGAGAAATTCAAATGCTATGACAAATGATCTCCCTGGA 2103
 Db 2041 GCTAATGGAGTGTGTTTGGAGAAATTCAAATGCTATGACAAATGATCTCCCTGGA 2100
 Qy 2104 CTATGCAAAATTTGGAAAAGCTGCTGGCAGAGGGTGTCTTCTAGGTTCAAGACACC 2163
 Db 2101 CTATGCAAAATTTGGAAAAGCTGCTGGCAGAGGGTGTCTTCTAGGTTCAAGACACC 2160
 Qy 2164 AAAGATTAATAATTTAACTCGGGGACCTACATATGATTCCTATAGTTTGAATACCTG 2223
 Db 2161 AAAGATTAATAATTTAACTCGGGGACCTACATATGATTCCTATAGTTTGAATACCTG 2220
 Qy 2224 GAAAGATGGAGTGTGAGGTTCTCTTACCTGCTGCAACTATGGCAAGATT 2283
 Db 2221 GAAAGATGGAGTGTGAGGTTCTCTTACCTGCTGCAACTATGGCAAGATT 2280
 Qy 2284 GAGCCGAGCATGTGAAGCTATGCTATGCAAGCATGCGAGAAAGTTTCTCCGCG 2343
 Db 2281 GAGCCGAGCATGTGAAGCTATGCTATGCAAGCATGCGAGAAAGTTTCTCCGCG 2340
 Qy 2344 TATACATATGAATCTGGGCTGAACCCAGAGATGCGAGAGAGTATAGTATGAT 2403
 Db 2341 TATACATATGAATCTGGGCTGAACCCAGAGATGCGAGAGAGTATAGTATGAT 2400
 Qy 2404 CTTGTTGGTAACTATGAGCCGATGATGAGCTGGTGTCTTATTTGCAAGAGCTGA 2463
 Db 2401 CTTGTTGGTAACTATGAGCCGATGATGAGCTGGTGTCTTATTTGCAAGAGCTGA 2460
 Qy 2464 AGACCTCTGAGAACTTTGAAACTATGATTAATGCAATTCGAGCTGGGCTCTCAGAGAGT 2523
 Db 2461 AGACCTCTGAGAACTTTGAAACTATGATTAATGCAATTCGAGCTGGGCTCTCAGAGAGT 2520
 Qy 2524 AGCCTTATGAACTATCTGTTGAAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGTG 2583
 Db 2521 AGCCTTATGAACTATCTGTTGAAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGTG 2580
 Qy 2584 AAGATCTAGCTGCTGCTGCTGCAATTCGATTCATGCTTCAAGCCAGAGATTTT 2643
 Db 2581 AAGATCTAGCTGCTGCTGCTGCAATTCGATTCATGCTTCAAGCCAGAGATTTT 2640
 Qy 2644 CTTAAAGCAGCTCATCTTTCAACGAGAGATATGTTTCTTATAGAAATCTGATGTC 2703
 Db 2641 CTTAAAGCAGCTCATCTTTCAACGAGAGATATGTTTCTTATAGAAATCTGATGTC 2700
 Qy 2704 GCTACCAATAGTATGATTAATGATGCAATTTTCAATATCTGCAATGCTCAAAATATGC 2763
 Db 2701 GCTACCAATAGTATGATTAATGATGCAATTTTCAATATCTGCAATGCTCAAAATATGC 2760
 Qy 2764 TTTGTTTGTAGCTAAGAACATAGTCCCATTAATACATGCTCCAAAGTTTATACCAAG 2823
 Db 2761 TTTGTTTGTAGCTAAGAACATAGTCCCATTAATACATGCTCCAAAGTTTATACCAAG 2820
 Qy 2824 ATTAAACAAGTTGCTAGATAATTTCACTAATATGCTGCTGAATTTTATGATCAAACTG 2883
 Db 2821 ATTAAACAAGTTGCTAGATAATTTCACTAATATGCTGCTGAATTTTATGATCAAACTG 2880

Qy 2884 TAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTGAATAGATAGATTAGAGA 2943
 Db 2881 TAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTGAATAGATAGATTAGAGA 2940
 Qy 2944 TTGCTTATGCTGCTTGTCTTCAACTTTTCTTCTGATTTTCTTTTCAATTAGG 3003
 Db 2941 TTGCTTATGCTGCTTGTCTTCAACTTTTCTTCTGATTTTCTTTTCAATTAGG 3000
 Qy 3004 GTCAGTACAGCTGACATTCAGAGACCTCCAGAAATGGAATGCTGAGCATGAGAGAA 3063
 Db 3001 GTCAGTACAGCTGACATTCAGAGACCTCCAGAAATGGAATGCTGAGCATGAGAGAA 3060
 Qy 3064 TATAGTATCCAAAGTGGCAGAAATTAAGTCTGCTTTTGGGCTGATCAAGCCATAGA 3123
 Db 3061 TATAGTATCCAAAGTGGCAGAAATTAAGTCTGCTTTTGGGCTGATCAAGCCATAGA 3120
 Qy 3124 AATGTTACAGAGGTAGAGGAATTAATCTACATTAATCAATGTTGTGAATCTGTTGG 3183
 Db 3121 AATGTTACAGAGGTAGAGGAATTAATCTACATTAATCAATGTTGTGAATCTGTTGG 3180
 Qy 3184 ACATGATTAATGCTGCTGCTTGTGATTCCTGTTATTAATAGTTTGAATGGGCGAA 3243
 Db 3181 ACATGATTAATGCTGCTGCTTGTGATTCCTGTTATTAATAGTTTGAATGGGCGAA 3240
 Qy 3244 TGTGAAGATTTGACATGACAGACAGCTGAACTGGCAGCTTGGGTTATGATT 3303
 Db 3241 TGTGAAGATTTGACATGACAGACAGCTGAACTGGCAGCTTGGGTTATGATT 3300
 Qy 3304 ATACACTGTTGAACTATCTGTTGACAGATGACAGTCTCAGACATAGAAACCCGTGCTC 3363
 Db 3301 ATACACTGTTGAACTATCTGTTGACAGATGACAGTCTCAGACATAGAAACCCGTGCTC 3360
 Qy 3364 TGTGGAAGCAACTCTGAGAGAGTCTGCTGTTATGATTTGATTTGATTTGATTTGATTT 3423
 Db 3361 TGTGGAAGCAACTCTGAGAGAGTCTGCTGTTATGATTTGATTTGATTTGATTTGATTT 3420
 Qy 3424 ATGCTACTGATGTCAGAACCTACACAAACAGATACGAATTTTCTGCTCAAGTCAAGGT 3483
 Db 3421 ATGCTACTGATGTCAGAACCTACACAAACAGATACGAATTTTCTGCTCAAGTCAAGGT 3480
 Qy 3484 GGAATATCACTGAAGGCTCTGTTCTTGCAATCAATATTAATCAATATGATGATGCTGAG 3543
 Db 3481 GGAATATCACTGAAGGCTCTGTTCTTGCAATCAATATTAATCAATATGATGATGCTGAG 3540
 Qy 3544 CTTGCGAGATTTCTTTGTTCTGTAATTCCTCTCTAAGTATGTTTAAATGAAACA 3603
 Db 3541 CTTGCGAGATTTCTTTGTTCTGTAATTCCTCTCTAAGTATGTTTAAATGAAACA 3600
 Qy 3604 CAAAAAATTAAGCTCTGTCGACACACCTTTTCTGATCTAATCTAATCAATTAAGGCT 3663
 Db 3601 CAAAAAATTAAGCTCTGTCGACACACCTTTTCTGATCTAATCTAATCAATTAAGGCT 3660
 Qy 3664 ACNA 3667
 Db 3661 ACNA 3664
 RESULT 2
 AY221467 3664 bp DNA linear PLN 05-AUG-2003
 LOCUS AY221467
 DEFINITION Arabidopsis thaliana truncated division protein (arce) gene, arce-1
 allele, complete cds, nuclear gene for chloroplast product.
 ACCESSION AY221467
 VERSION AY221467.1 GI:33436274
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 3664)
 REFERENCE Vitsha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Eyrp,H. and
 AUTHORS

Oteyyoung, K.W.
ABC6 is a J-Domain Plastid Division Protein and an Evolutionary
Descendant of the Cyanobacterial Cell Division Protein Ftn2
JOURNAL
PUBMED
Plant Cell 15 (8), 1918-1933 (2003)
12897262
2 (bases 1 to 3664)
Vitha, S., Koksharova, O., van Erp, H., Froehlich, J.E. and
Oteyyoung, K.W.
Direct Submission
Submitted (17-JAN-2003) Department of Plant Biology, Michigan State
University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
Source
1..3664
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/ecotype="Ws-2"
/note="genotype: arc6 mutant"
1..3664
/gene="arc6"
/allele="arc6-1"
join(<478..984,1067..1291,1379..>1621)
/gene="arc6"
/allele="arc6-1"
/product="truncated division protein"
join(478..984,1067..1291,1379..1621)
/gene="arc6"
/allele="arc6-1"
/note="truncation caused by premature stop codon"
/codon_start=1
/product="truncated division protein"
/protein_id="AAQ18644.1"
/db_xref="GI:33436275"
/translation="MEALSHVIGLSLPOLCRLEPATTKLRBSHNTSTTCSAKMWD
RLSDPRPTDSSSSPRATRTATRTVLSPLSDPRRHPIPIEDYQVGAQTHPT
DGRRAEARVSKPQGFSDDLISRQILQACETLSNPSRRKRNESGLDDEEAT
VITDVPWIKVGA/CVLQEGEETIVRVEBALKEPLPSFKQDVLA/VMAFLDVS
RDMALDLPDITGYEFVEBALKILQEGSISLAPDLRAIDITLEETITPRYVELLIG
LPLGDVAAKRLNGLSGVRNLIWSVGGSALVGLTRKFMNEARL"

ORIGIN
Query Match 99.8%; Score 3660.8; DB 8; Length 3664;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3662; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TCTGCATTAGAGAAATACAAATTATTAAGCAATTGCTGATTTCAACAAGATTGCTT 63
DB 1 TCTGCATTAGAGAAATACAAATTATTAAGCAATTGCTGATTTCAACAAGATTGCTT 60
QY 64 GGCATATAGATTCAATTGGCTCTGTTTGCCTTTACATTTACATGTCATATAGTTGCAAT 123
DB 61 GGCATATAGATTCAATTGGCTCTGTTTGCCTTTACATTTACATGTCATATAGTTGCAAT 120
QY 124 TTTAACAATTCAGTTGAGATGTTAAGAAAGAGAGCAATTGATGGGTTTGTGGCTT 183
DB 121 TTTAACAATTCAGTTGAGATGTTAAGAAAGAGAGCAATTGATGGGTTTGTGGCTT 180
QY 184 AAACCTTAAGTGTCAAGAAATTAAGTCATTTGTTTCTGTGCTCAATAGTGAAT 243
DB 181 AAACCTTAAGTGTCAAGAAATTAAGTCATTTGTTTCTGTGCTCAATAGTGAAT 240
QY 244 GAAGGCAACTCAACGGTTCTTAGTGAATAGATTATTTAGACATTTAAACATCAATAA 303
DB 241 GAAGGCAACTCAACGGTTCTTAGTGAATAGATTATTTAGACATTTAAACATCAATAA 300
QY 304 GTCCGTGGCGACGTGAACATCAATGATGTTTTTTTATTTTTCAGTAGCTGTGATGTT 363
DB 301 GTCCGTGGCGACGTGAACATCAATGATGTTTTTTTATTTTTCAGTAGCTGTGATGTT 360
QY 364 TTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 423
DB 361 TTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420

QY 424 TCTCTCAATATGTAAACAAACAAAGTTTGTGTAGTACCTTAACAAACATCCCATG 483
DB 421 TCTCTCAATATGTAAACAAACAAAGTTTGTGTAGTACCTTAACAAACATCCCATG 480
QY 484 GAAGCTGTGATGACGTGGCGCATTTGGTCTCTCCCATTTCAATTAATGCGATTACACCG 543
DB 481 GAAGCTGTGATGACGTGGCGCATTTGGTCTCTCCCATTTCAATTAATGCGATTACACCG 540
QY 544 GCGACGACAAAGCTCCGACGTAGCCACACACACCTTAACAATCTTGTCTCGCGACGAAA 603
DB 541 GCGACGACAAAGCTCCGACGTAGCCACACACCTTAACAATCTTGTCTCGCGACGAAA 600
QY 604 TGGGCGGACCGTCTTCTCTCCGACTTCAATTTCACTTCGATTCCTCTCTCTCTTCC 663
DB 601 TGGGCGGACCGTCTTCTCTCCGACTTCAATTTCACTTCGATTCCTCTCTCTCTTCC 660
QY 664 GCCACGCGCACACACACCGGCACTCTGTCCTGACACCATCTATTGATCGTCCGAA 723
DB 661 GCCACGCGCACACACACCGGCACTCTGTCCTGACACCATCTATTGATCGTCCGAA 720
QY 724 CGCCACGTCCCAATCCCATTTGATTTCTACAGATATTAGAGCTCAACACATTTCTTA 783
DB 721 CGCCACGTCCCAATCCCATTTGATTTCTACAGATATTAGAGCTCAACACATTTCTTA 780
QY 784 ACCGATGGAATCAAGAGCAATTCGAAGCTTAGGTTTCGAACCGCGCAATTCGGTTTC 843
DB 781 ACCGATGGAATCAAGAGCAATTCGAAGCTTAGGTTTCGAACCGCGCAATTCGGTTTC 840
QY 844 AGGACGACGCTTTAATCAGCCGCGACACAGATTTCTTCAAGCTGTCGGAACCTGTCT 903
DB 841 AGGACGACGCTTTAATCAGCCGCGACACAGATTTCTTCAAGCTGTCGGAACCTGTCT 900
QY 904 AATCTCGGCTGTAAGAGAGTACAAATGAAGTCTTCTTGATGATGAAGAGCTACAGTC 963
DB 901 AATCTCGGCTGTAAGAGAGTACAAATGAAGTCTTCTTGATGATGAAGAGCTACAGTC 960
QY 964 ATCACTGATGTTCTTGGATTAAGTGAATTCGATTTCCGAATTAATGAATTTCTTCTT 1023
DB 961 ATCACTGATGTTCTTGGATTAAGTGAATTCGATTTCCGAATTAATGAATTTCTTCTT 1020
QY 1024 TTAAATTCAGAAATGGAATTAAGAAAGAACTTTTATCTAGTGAAGTCTCTGGGCTCT 1083
DB 1021 TTAAATTCAGAAATGGAATTAAGAAAGAACTTTTATCTAGTGAAGTCTCTGGGCTCT 1080
QY 1084 CTGTGATTCGAAGAGTGTGAGACTGAGATAGTCTTCTCGGTTGTGTAGGCTCTGCT 1143
DB 1081 CTGTGATTCGAAGAGTGTGAGACTGAGATAGTCTTCTCGGTTGTGTAGGCTCTGCT 1140
QY 1144 TAAAGAGAGTCTTAAGTCTGTTTAAAGCAAGTGTGTTTAAATGAGCGCTTGCGTT 1203
DB 1141 TAAAGAGAGTCTTAAGTCTGTTTAAAGCAAGTGTGTTTAAATGAGCGCTTGCGTT 1200
QY 1204 TCTCGATGTCGAGGAGTCTATGTCATTTGCAACCTGATTTTATTACTGTTATGA 1263
DB 1201 TCTCGATGTCGAGGAGTCTATGTCATTTGCAACCTGATTTTATTACTGTTATGA 1260
QY 1264 GTTGTGTAAGAGCTTTGAAGCTTTTACAGATGTTGACTTGTGTTAATTTGACG 1323
DB 1261 GTTGTGTAAGAGCTTTGAAGCTTTTACAGATGTTGACTTGTGTTAATTTGACG 1320
QY 1324 AGCGTGGCTTTAAGAACTTTCTTGATTTGAATCTTGTATTAAGTCTGTGTAGGA 1383
DB 1321 AGCGTGGCTTTAAGAACTTTCTTGATTTGAATCTTGTATTAAGTCTGTGTAGGA 1380
QY 1384 GGAAGGACAAAGTACGCTTGACACCGAATTAAGTGAACAAATTAATGAAGACTTTGAAAG 1443
DB 1381 GGAAGGACAAAGTACGCTTGACACCGAATTAAGTGAACAAATTAATGAAGACTTTGAAAG 1440
QY 1444 GATCACTCGCGGTTATGTTGTGAGGCTAATGAGCTTAACGCTTGATGATTAACGCTGC 1503
DB 1441 GATCACTCGCGGTTATGTTGTGAGGCTAATGAGCTTAACGCTTGATGATTAACGCTGC 1500
QY 1504 GAAAGACTAAATGTTTAAAGCGGTGTGCGGAATATTTGTGTGTGTGTGAGAGAGTGG 1563

[illegible]

Db	2581	AAGATCTAGCTGCTGGTGTGGCAATTGGA	CTGATTTCACTGTTCAAGCAAGATATTTT	2640
Qy	2644	CTTAAGAAGCACTCATCTTTTCAACGCAAG	ATAGTGTTCCTTATAGAAATCTGATGTC	2703
Db	2641	CTTAAGAAGCACTCATCTTTTCAACGCAAG	ATATGTTCTTCTATAGAAATCTGATGTC	2700
Qy	2704	GCTACCAATAGTATGATTTAAATGATGAA	TTTTCATATCTGATGCTCAAAATATGCG	2763
Db	2701	GCTACCAATAGTATGATTTAAATGATGAA	TTTTCATATCTGATGCTCAAAATATGCG	2760
Qy	2764	TTGTTTTGTAGCTAAGAACATAGTTCCTCA	CTTATATACATGTCCTCCAAAGTTGTACAG	2823
Db	2761	TTGTTTTGTAGCTAAGAACATATTTCCACT	TAATATACATGTCCTCCAAAGTTGTACAG	2820
Qy	2824	ATTAACAAGTTGCTGAGTAAATTTCACTAA	TATATGCTGCTGGAATTTTTCATCAAC	2883
Db	2821	ATTAACAAGTTGCTGAGTAAATTTCACTAA	TATATGCTGCTGGAATTTTTCATCAAC	2880
Qy	2884	TAGACAGAAATGTAAATTTTCACTCTCAAC	ATTTCTGTTTGAATTAAGCTAGATTAAGA	2943
Db	2881	TAGACAGAAATGTAAATTTTCACTCTCAAC	ATTTCTGTTTGAATTAAGCTAGATTAAGA	2940
Qy	2944	TTGCTTAGTGTGCTTTGTCCAACTTTTCTT	CTTGATTTTTCCTTTTCGATTTTAGG	3003
Db	2941	TTGCTTAGTGTGCTTTGTCCAACTTTTCTT	CTTGATTTTTCCTTTTCGATTTTAGG	3000
Qy	3004	GTCAGTCAGAGCTGACGATTCAGAGA	CTTCCAGAAATGAGATCTAGAGCTGACAGAA	3063
Db	3001	GTCAGTCAGAGCTGACGATTCAGAGA	CTTCCAGAAATGAGATCTAGAGCTGACAGAA	3060
Qy	3064	TATAGTATCCAGATGCGAGAAATTAAG	CTCTGCGCTTTTGGGCTGTATCACCGCATAGA	3123
Db	3061	TATAGTATCCAGATGCGAGAAATTAAG	CTCTGCGCTTTTGGGCTGTATCACCGCATAGA	3120
Qy	3124	AATGTTACCAAGGATGAGGAAATTAATCTA	CAATTCATCAATTTGTGTGAAAACCTGTGG	3183
Db	3121	AATGTTACCAAGGATGAGGAAATTAATCTA	CAATTCATCAATTTGTGTGAAAACCTGTGG	3180
Qy	3184	ACATGATTAATAGTCTGGTGCCTTGTGAT	TTTATAGATTTTGGATGGGCGAA	3243
Db	3181	ACATGATTAATAGTCTGGTGCCTTGTGAT	TTTATAGATTTTGGATGGGCGAA	3240
Qy	3244	TGCTGAAATTTGGACTGACAGAGCAGTGA	AACTGGGCACTGGGTGGTTATGATT	3303
Db	3241	TGCTGAAATTTGGACTGACAGAGCAGTGA	AACTGGGCACTGGGTGGTTATGATT	3300
Qy	3304	ATPACAGTTGAAATCTATCTGTGACAGTGA	CAGTCTGACGATGGAACCCGTGCTC	3363
Db	3301	ATPACAGTTGAAATCTATCTGTGACAGTGA	CAGTCTGACGATGGAACCCGTGCTC	3360
Qy	3364	TGTTGGAAGCAACTCTGGAGAGTCTGGTTC	CTATCTGATTTGGTTCAATCAGAAAAACA	3423
Db	3361	TGTTGGAAGCAACTCTGGAGAGTCTGGTTC	CTATCTGATTTGGTTCAATCAGAAAAACA	3420
Qy	3424	ATGCTACTGATGTCAGAACTTACACAA	CAGATACGAAATTTTCTGGTCCAACTCAGGCT	3483
Db	3421	ATGCTACTGATGTCAGAACTTACACAA	CAGATACGAAATTTTCTGGTCCAACTCAGGCT	3480
Qy	3484	GGAATAATCAGGAAGCTCTGTCTGTGAT	ATATATATCTCATATGATGAGATGCTGAG	3543
Db	3481	GGAATAATCAGGAAGCTCTGTCTGTGAT	ATATATATCTCATATGATGAGATGCTGAG	3540
Qy	3544	CTTGGAGATTTCTCTTGTCTGTAAATCT	CTCTCTAAGTTAGTGTATTAATATGAACA	3603
Db	3541	CTTGGAGATTTCTCTTGTCTGTAAATCT	CTCTCTAAGTTAGTGTATTAATATGAACA	3600
Qy	3604	CAAAAAATTAACGTTCTTGGCAAC	CCCTTTTCTTGTATCTAAATCTATTAAGGCT	3663
Db	3601	CAAAAAATTAACGTTCTTGGCAAC	CCCTTTTCTTGTATCTAAATCTATTAAGGCT	3660
Qy	3664	ACAA 3667		
Db	3661	ACAA 3664		

RESULT 3
LOCUS AB016888
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MDH9.
ACCESSION AB016888 BA000015
VERSION AB016888.1 GI:3449329
SOURCE
ORGANISM Arabidopsis thaliana (chale cress)
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.
1
Asamizu,E., Sato,S., Kaneko,T., Nakamura,Y., Kotani,H., Miyajima,N. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 5. VIII. Sequence features of the regions of 1,081,958 bp covered by seventeen physically assigned P1 and YAC clones
DNA Res. 5 (6), 379-391 (1998)
99156233
PUBMED 10048488
2 (bases 1 to 85791)
Nakamura,Y.
Direct Submission
Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:yamakum@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)
Addresses for correspondence: ka@kazusa.or.jp
For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MDH9
Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://complib.ornl.gov/Grail-1.3/>), GENSCAN (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://gremli.ni.zool.iastate.edu/cgi-bin/bp.cgi>).
Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).
This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions.
The 5' clone is K5J14 and the 3' clone is K16E1.
Location/Qualifiers
1. 85791
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MDH9"
/clone_1lb="Mitau1 P1"
/ecotype="Columbia"
/complement(59..239)
/note="CDS is reported in Acc# AB023032
gene_id:K5J14.12
unknown protein"
/number=1
/evidence=not_experimental
/complement(join(2056..2079,2278..2393,2991..3633))
/note="contains similarity to unknown protein
gene_id:MDH9.1
p1c|T04864"
/pseudo
/codon_start=1
/evidence=not_experimental
join(5739..6138,6462..6595)

CDS
/note="unnamed protein product; gene_id:MDH9.2
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BABI0474.1"
/db_xref="GI:9759472"
/translation="MWKDEFAEDDRKSSNLAKNSVPLPERGRFYEEYAKNE
RLRRKGEEAVGEGVAKGTPIYNGVNHPEPTNRGRGAKESIKIKTVASVSSSIPR
YSLSRAKERNRPPPLPNAVASAMKSVTRKEERLSLLCTACNTSSGLAIFRLTGS
EPGVHKRRTLPSPALV"
complement(join(6868..7911,7989..8414,8555..8671,
8754..9137))
/note="gene_id:MDH9.3"
/codon_start=1
/evidence=not_experimental
/product="am repeat containing protein"
/protein_id="BABI0475.1"
/db_xref="GI:9759470"
/translation="WDTEPATGDAENRDEBTAEPIHDEVYDAVEIHEEYKEDD
DCGLVSDIVSVEFLDQINGYRRTQKCEFNIVRLKILIPFLDEIRGEFSCKHF
LNLRLRVFLAAKLTLECSNGSKTYMALDEETMTFRHSIYEKLNELVAVAPDEMI
SGADKBIDSLCKQLKAKRRDTODIELADMMVVSFSDPDPNADSAIETLAKLE
LQTTDLKTEITAIQSLIQDKGLNIEIKOHIIEILNKFKQLGEATDILYQVINK
AATKSTSLIPHEFLCITTEINLDPVLIATGOTYKEKSTOKPFDGHTCKPTROEL
DHLSLAPNPLAKNLIMQCEKKNFKIPEKVSFDSQEQEDVSLVEALSSQLEEQ
RSVVKQMRLLARENENRVLINAVGAIPLVLQSLYPSGQIQNAVVTLLNLSIDEN
KKLSISGALPNIIEILENGNREARERNSAALFSLMLDENKTYIGLSNGIPLVLDL
OHTLNGKQALPALFNLISANKRKAIDAGIVQPLNLNKKNGKIMDEALSLIL
LASHPGROALIGQLSFEITLVEFTRGTPNKCAVSLLELSSNNSSFLAALQGV
YRIIVETTSNGTRAKRNALIQLSKSGQT"
10888..12579
/note="unnamed protein product; gb|AAF07802.1
gene_id:MDH9.4
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BABI0476.1"
/db_xref="GI:9759471"
/translation="MMISEKPLGVEESIRDLLEVLTVSRRLVKSQKTKKKIKHTEV
VEBEIVARGAVNCLISVGCRAVDTEGDEPDSSNKMSSASEGKGLMTICGEETRL
DCGSYVRRBPRFKNNRKYLADSGOYRKHVYLPDILMLCMRLPLTSLNHLVCK
KMSMANQRFLOMRREGSGQTPTWLPFLPALDGCSSGDIHGVDQODKMRILETLL
KGFMTSVSIHBEIITVIGRSMDRNSFSSHRIIVFSPITAKRKIASRHKASLPY
VGAETVTSSEPTQTKQNRQDRPHLSRVGSESDVDEPHRLSVRQNRNSADQNGTK
SHRLIQKLDRLNRNSKRFVLIAGTGLFDEPLDSGEIYDSATVTSWEMQRLPMGF
GVVSGCIIINGIIFYAASENDKSLSGYIEIGGFITITQSPILPRAVHEFYPLVSCNRL
FMLSVCBQBGQIGRRNKAVKMELDLVYITTEVSVHPDAPMDMATVYSDNI
LMKIEPMFKIFGYLSFFTYCDVUTTEASRHRHSRNRSSQNLNCTNKITIALHLT"
complement(15087..16778)
/note="unnamed protein product; gb|AAF07802.1
gene_id:MDH9.5
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BABI0477.1"
/db_xref="GI:9759472"
/translation="MMISEKPLGVEESIRDLLEVLTVSRRLVKSQKTKKKIKHTEV
EDEVARGAVNCLISVGCRAVDTEGDEPDSSNKMSSASEGKGLMTICGEETRL
DCGSYVRRBPRFKNNRKYLADSGOYRKHVYLPDILMLCMRLPLTSLNHLVCK
KMSMANQRFLOMRREGSGQTPTWLPFLPALDGCSSGDIHGVDQODKMRILETLL
KGFMTSVSIHBEIITVIGRSMDRNSFSSHRIIVFSPITAKRKIASRHKASLPY
VGAETVTSSEPTQTKQNRQDRPHLSRVGSESDVDEPHRLSVRQNRNSADQNGTK
SHRLIQKLDRLNRNSKRFVLIAGTGLFDEPLDSGEIYDSATVTSWEMQRLPMGF
GVVSGCIIINGIIFYAASENDKSLSGYIEIGGFITITQSPILPRAVHEFYPLVSCNRL
FMLSVCBQBGQIGRRNKAVKMELDLVYITTEVSVHPDAPMDMATVYSDNI
LMKIEPMFKIFGYLSFFTYCDVUTTEASRHRHSRNRQNLNCTNKITIALHLT"
complement(join(18111..18193,18281..18535,18773..19001),
19088..19195,19307..19447,19542..19952,20129..20245))
/note="unnamed protein product; gene_id:MDH9.6
similar to unknown protein
bp|P42251"
/codon_start=1
/evidence=not_experimental

Query Match 99.8%; Score 3659; DB 8; Length 85791;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3662; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

CDS

1 TGTTCGATTAAGAGAATACATTAATTAAGCAATTTGCTTGAATTCACAGATTTTG 60
 Db 63597 TGTTCGATTAAGAGAATACATTAATTAAGCAATTTGCTTGAATTCACAGATTTTG 63656
 Oy 61 CTTCGCTATAGGATTCATTTGCTCTGTTGCTTTTACATTTACATGTCATATAGTTTCG 120
 Db 63657 CTTCGCTATAGGATTCATTTGCTCTGTTGCTTTTACATTTACATGTCATATAGTTTCG 63716
 Oy 121 AATTTTACATTTCAATTTGATTTAAGAAAAGAGCAATTTGATGGCGTTTGTGCGG 180

CDS

gene_id="BAB10478.1"
 /db_xref="GI:9759473"
 /translation="MNQISFWCLSLTLTVVIAAESTPTITRLAFSGCANQASAPQPIWDA
 INKDPQLPFLMAGDNIYGDIRRLPLVIGERTFPGWNRSPRPVPSSEEMKLRKAKA
 ANPGYSRIQRKAKVIGTDHYDGLNDAGKEFDRIYINQKMLDELNEPLSPRRKA
 GYVAATYTPGPRKRYKVIYLDTRYHRDPLRSGSLIGTDQMDNEISGPRSETII
 GSVAYTINISATKGVPLFMYSWCKGPPKERRRLPCLIDTQDQVYIPISGVHREIT
 RYDCSVGRLPDTVTSGLVQSVKRYVPPPLRSITVRLTWYIPSTIRVINDCKPKSCT
 YGQNFEGAIISIDWNAIPVITIKLIRIDVAGDVLVGNVLSLQPGSSNSLKDQTTKQ
 SQRYCTLBLELPGTIRYALVLIPTFLIAMLAIIGALVATLACVYCKKD"
 complement(21540..22097)
 /note="unnamed protein product; contains similarity to
 calmodulin
 gene_id:MDH9.7"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB10479.1"
 /db_xref="GI:9759474"
 /translation="WTLLAKNQSSLSRLYKRYKSSRRSSRLREBESSTNSGSSS
 LNVNELRFDMDNSDKISGEELQCVSLIGDALSREVEEVKISDVGDGFDID
 FEEELKGEGBGSDSEBRKELKRAFGYVMEGSEPTITASRLTSLRSGSCVTYDAC
 KMIKRGPNDDGVLSFDEPFLMKR"
 join(24090..24288,24410..24614,24701..24945,25115..25313,
 25632..25679,25821..25928,26039..26227,26314..26613,
 26730..26831,27169..27303,27487..27711,27919..28122,
 28213..28380,28646..28795,28889..29026,29204..29431,
 29519..29671,29784..29910,30115..30241,30554..30674,
 30763..30843,30932..31033,31112..31262,31338..31429)
 /note="gene_id:MDH9.8"
 /codon_start=1
 /evidence=not experimental
 /product="pitr1lyain"
 /protein_id="BAB10480.1"
 /db_xref="GI:9759475"
 /translation="MASSSSSIFGVKSPITLAPRSGSDSRRLYLRKMKVFNPS
 SPRLPHVRVYRABSLIPYMLGLMAOPSHGRCRMTVSGKEAGNISGGRPCLT
 CKRNOAGIRRALPSAFVDRTRAFSLRSSLTSLRKSQIVNATLQDPEPHAGTRMPD
 GIVARQDLDLPEIDSLAEALFGLCPSPHKLPRQLNQKARNATLLPKVPPNRP
 BAHMEVHDSIDBEDBQGLAHMIEHVAFLPSKREKELGTGARSNAITLDRPHVFIH
 HSPTTKQSEDDLPSVLDALNEIHFHFKPLSRVEKERRAILSELQMMNTIERYVDC
 OLQHLHSENGLGRFPFGLSEBOIKKMDVDKIKRHEHWFPANATLYIVGDIINIPR
 IVHNIEAVFGKGLDNESTPSPRPGAGMANLPVPLPGGLGTFSENEKTIOTR
 KMTIRBRAIRPVEHMSLPETSVDLKPQIFKHELLQNRALNPKCIPIKSIOTRG
 DLRLVLMKRIPLSALHFRINTRYKSNPPTFVSLDHSRREGCTVITTLVLEPQN
 WQNAKVAVQVEGRRLKEGVTRGELTRTMDLKDSEHMAINDVSEVDMLDEPMS
 DALSTWDOQGHETLVAAGVTLEEVNTGAVALFEIIDPGFTRPLPAALYACV
 PKTVHVDGSGSDENIISPDIEIISVKSGLPIABEPELVPEKELISQSLQELTLOR
 MPCVPPIPGSGCTKLHDKETCIQURLSNGVIANYKSTSPSRAGVALLYGGRRAL
 TSDSGAVAVGRTLSBEGVDFRREOVLEFCVNLINCSLSRESEETFLAEFPFTR
 DNGVQAPQLHMLYERSVLEDAEDRARQOLYAFSRNIPSLSEALTKIMILMNG
 DERFEAPQLHMLYERSVLEDAEDRARQOLYAFSRNIPSLSEALTKIMILMNG
 ASHSGAPRGESEPIFRQPTAGLOEQVFLKQDTERAPCIYAGPAGNMGVGTGODL
 FOSVSKIPVAHDGILKSEREOLRGGEDRLQKTRAPLPCVYPMGLAIETINSLFTT
 VRDSIGLTYDSFEINLEPDLRLNTGTYIVSTSTPKRYKADACKSVLRGHSNQIAP
 RELDPAKRTLLRHEBARLKSNAVWLTNLALHQAQSVPRKEELISKEIVLSEASIED
 IYLVNQLRVDESDLSYSCIGIAGAQAGEIRITVLSBEBEPDFSGVAVPVGSGSWTR
 PTT"
 complement(join(33251..33326,33512..33663,33744..33797,

Db	63717	AAATTTCACATTTCACTTCAGTTGCAATGTTAAAGAAAGAGGAAATTAATGAGGCTTTGTGGG	63776
QY	181	TTTAACTTTAAAGTAGTCAAGATTAAAGTCAATTGGTTTACTGTGCTCTAATATGTGTAA	240
Db	63777	TTTAAACTTTAAAGTAGTCAAGATTAAAGTCAATTGGTTTACTGTGCTCTAATATGTGTAA	63836
QY	241	AAATCAAGCAATCCCAACGGTTCTTAAAGTGCAGATTAATTAATTAAGACATTTAAACATCAT	300
Db	63837	AAATCAAGCAATCCCAACGGTTCTTAAAGTGCAGATTAATTAATTAATTAAGACATTTAAACATCAT	63896
QY	301	AAAGTCGTGGCGCACTGTAACTATAGATTTGTTTTTATTTTTCATAGAGCTGTGAT	360
Db	63897	AAAGTCGTGGCGCACTGTAACTATAGATTTGTTTTTATTTTTCATAGAGCTGTGAT	63956
QY	361	GTTTTTGTATTTAATTAACTTAATCTACATAATTCAAAATTCATTAACCTTAGACGACCAA	420
Db	63957	GTTTTTGTATTTAATTAACTTAATCTACATAATTCAAAATTCATTAACCTTAGACGACCAA	64016
QY	421	CAGTCTCTTCAATATGTAAACAAGAACAAAGTTTTGTATAGTACCTTAAATAACATCTCCC	480
Db	64017	CAGTCTCTTCAATATGTAAACAAGAACAAAGTTTTGTATAGTACCTTAAATAACATCTCCC	64076
QY	481	ATGGAAGCTGTAGTACGTCGCGGCAATTTGCTCTCCCAATTCGAATTAAGCGATTACCA	540
Db	64077	ATGGAAGCTGTAGTACGTCGCGGCAATTTGCTCTCCCAATTCGAATTAAGCGATTACCA	64136
QY	541	CCGCGCAACGACAAAGCTCCGACGTAGGCAACACACTCTAACAATCTGTCTCCGCAAC	600
Db	64137	CCGCGCAACGACAAAGCTCCGACGTAGGCAACACACTCTAACAATCTGTCTCCGCAAC	64196
QY	601	AAATGGGCGCAACGTCCTTCTCTCCGACTTCAATTTCAACTCCGATTCCTCTCTCC	660
Db	64197	AAATGGGCGCAACGTCCTTCTCTCCGACTTCAATTTCAACTCCGATTCCTCTCTCC	64256
QY	661	TTGGCCACCGCCACACACACGCGCACTCTGTCTCTGCGACCACTATATGATGCTGCC	720
Db	64257	TTGGCCACCGCCACACACACGCGCACTCTGTCTCTGCGACCACTATATGATGCTGCC	64316
QY	721	GAAAGCCACGTCGCCCATCCCAATGTAATTTCTACAGAGTATTAAGACCTCAAAACAATTTCT	780
Db	64317	GAAAGCCACGTCGCCCATCCCAATGTAATTTCTACAGAGTATTAAGACCTCAAAACAATTTCT	64376
QY	781	TTTAAACGATGGAATCAGAAAGACATTCGAAAGCTAAGGTTTCGAAACCGCGCAATTCGGT	840
Db	64377	TTTAAACGATGGAATCAGAAAGACATTCGAAAGCTAAGGTTTCGAAACCGCGCAATTCGGT	64436
QY	841	TTTACGCGACGACGCTTTAATCAGCCGCGAGACGATTTCTCAAGGCTGCGGAAACTCTG	900
Db	64437	TTTACGCGACGACGCTTTAATCAGCCGCGAGACGATTTCTCAAGGCTGCGGAAACTCTG	64496
QY	901	TCTTAATCCTCGGTCCTAGAAAGAGTACCAATGAAGGCTCTTCTGATGATGAAGAAGCTACA	960
Db	64497	TCTTAATCCTCGGTCCTAGAAAGAGTACCAATGAAGGCTCTTCTGATGATGAAGAAGCTACA	64556
QY	961	GTCATCACTGATGTTCTTGGGATTAAGTAAATTTGATTTCCGAAATTAATAAGTTCTTTC	1020
Db	64557	GTCATCACTGATGTTCTTGGGATTAAGTAAATTTGATTTCCGAAATTAATAAGTTCTTTC	64616
QY	1021	GTTTAAATTTCAAGAAATGGAATPAAGAAAGAACTTTAATCTAAGAAAGCTTCGCGGCT	1080
Db	64617	GTTTAAATTTCAAGAAATGGAATPAAGAAAGAACTTTAATCTAAGAAAGCTTCGCGGCT	64676
QY	1081	TCTCTGTGTATTCAGAAAGGTTGTGAGACTGAGATGATTTCTTCCGGTGTGTGAGGCTCT	1140
Db	64677	TCTCTGTGTATTCAGAAAGGTTGTGAGACTGAGATGATTTCTTCCGGTGTGTGAGGCTCT	64736
QY	1141	GCTTAAAGAGAGGTTGCTTAAGTGTGTTTAAAGCAAGTGTGTTTAAAGTAAAGGCTTTC	1200
Db	64737	GCTTAAAGAGAGGTTGCTTAAGTGTGTTTAAAGCAAGTGTGTTTAAAGTAAAGGCTTTC	64796
QY	1201	GTTTTCGATATCTCGAAGGATTCGATATGCAATTTGATTCACCTGATTTATTAATCTGTATA	1260
Db	64797	GTTTTCGATATCTCGAAGGATTCGATATGCAATTTGATTTCAACCTGATTTATTAATCTGTATA	64856

OY	1261	TGAGTTGTTGAGGACGTTTGAAGCTTTACAGGTATGACGCTTGGTAAATTG	1320
Db	64857	TGAGTTGTTGAGGACGTTTGAAGCTTTACAGGTATGACGCTTGGTAAATTG	64916
OY	1321	ACGAGCGTTGGCTTTATPAGAAGCTTTCTGTATTTGATCTTGTATTTAGATCTTGTA	1380
Db	64917	ACGAGCGTTGGCTTTATPAGAAGCTTTCTGTATTTGATCTTGTATTTAGATCTTGTA	64976
OY	1381	GGAGGAAGAGCAAGTACCTTGGACCGGATTTACGTGCACAAATTGATGAGACTTTGGA	1440
Db	64977	GGAGGAAGAGCAAGTACCTTGGACCGGATTTACGTGCACAAATTGATGAGACTTTGGA	65036
OY	1441	AGAGATCACTCCGGGTTATGTCCTTGGAGGCACTTGGCTTACGCGCTGGATGATTAAGC	1500
Db	65037	AGAGATCACTCCGGGTTATGTCCTTGGAGGCACTTGGCTTACGCGCTGGATGATTAAGC	65096
OY	1501	TGCGAAAAAGCTAATATGGTTTAAAGCGGTGCGGAATATTTGTCGTCTGTGGAGAGG	1560
Db	65097	TGCGAAAAAGCTAATATGGTTTAAAGCGGTGCGGAATATTTGTCGTCTGTGGAGAGG	65156
OY	1561	TGAGCATCAGCTCTTGTGGGGGTTGACCCGTGAGAGTTTATGATGAGCGCTTTT	1620
Db	65157	TGAGCATCAGCTCTTGTGGGGGTTGACCCGTGAGAGTTTATGATGAGCGCTTTT	65216
OY	1621	ACGATGACAGCTGCTGAGCAGGTATACGTTTATAGATACCTTTTTTAAATTTCTTATGCA	1680
Db	65217	ACGATGACAGCTGCTGAGCAGGTATACGTTTATAGATACCTTTTTTAAATTTCTTATGCA	65276
OY	1681	TGATATACCTTTAGGTTTCATATTTAATGATGTGTGTAGGTGATCTTTTGTGA	1740
Db	65277	TGATATACCTTTAGGTTTCATATTTAATGATGTGTGTAGGTGATCTTTTGTGA	65336
OY	1741	GCTACCCGACGCAATTTTCCAGCAGACTCTTGAAGTTTACGAATTTGCACTTGTCTT	1800
Db	65337	GCTACCCGACGCAATTTTCCAGCAGACTCTTGAAGTTTACGAATTTGCACTTGTCTT	65396
OY	1801	GTGGCTCAAGCTTTTATTTGGTAAAGCCACACTTTTACAGAGCTGTATTAAGCAATTC	1860
Db	65397	GTGGCTCAAGCTTTTATTTGGTAAAGCCACACTTTTACAGAGCTGTATTAAGCAATTC	65456
OY	1861	CAGCAACTTCAGCAGGCTAAGGTATGGCTATGAGATTCCTGCGATGTTATGATACA	1920
Db	65457	CAGCAACTTCAGCAGGCTAAGGTATGGCTATGAGATTCCTGCGATGTTATGATACA	65516
OY	1921	CGAATTAATTGGGAGATGAGACTTGGGTCTAGAAAGGGAATCTGTGCACTGCTTATAGCC	1980
Db	65517	CGAATTAATTGGGAGATGAGACTTGGGTCTAGAAAGGGAATCTGTGCACTGCTTATAGCC	65576
OY	1981	AAAGTGAATGAATCCGATATGTGGTTGGGCTTTAGACAGTGAAGATTCACAATATAGGAAT	2040
Db	65577	AAAGTGAATGAATCCGATATGTGGTTGGGCTTTAGACAGTGAAGATTCACAATATAGGAAT	65636
OY	2041	CCAGCTATTGTGAGATTGTTTGGAGAAATTCAAATCTGATGACATGATGATCTCCCT	2100
Db	65637	CCAGCTATTGTGAGATTGTTTGGAGAAATTCAAATCTGATGACATGATGATCTCCCT	65696
OY	2101	GGACTATGCAAAATTGTTGAAAATCTGGTTGGCAGGGGTTGTCTTTCTTAGGTTCAAGAC	2160
Db	65697	GGACTATGCAAAATTGTTGAAAATCTGGTTGGCAGGGGTTGTCTTTCTTAGGTTCAAGAC	65756
OY	2161	ACCAAAAGTAAAAAATTTTAACTCGGGGACCAATCATATGATGATGATCTATGTTTGAAGTAC	2220
Db	65757	ACCAAAAGTAAAAAATTTTAACTCGGGGACCAATCATATGATGATGATCTATGTTTGAAGTAC	65816
OY	2221	TTGGAAGAAGTGAAGTATGTTCAAGGTTCTCTTTAGCTGCTGTCGAATATGCGAAG	2280
Db	65817	TTGGAAGAAGTGAAGTATGTTCAAGGTTCTCTTTAGCTGCTGTCGAATATGCGAAG	65876
OY	2281	ATTGGAAGCCGAGCATGTGAAAAGCTAAGTCTATGCAAGGCACTGCAGAAAAGTTTTCTTCC	2340
Db	65877	ATTGGAAGCCGAGCATGTGAAAAGCTAAGTCTATGCAAGGCACTGCAGAAAAGTTTTCTTCC	65936

QY	2241	CGCATATCAATATGAAACTGGCTGAAACCCAGAGATGTCAGAGACATGTTTAGTGA	2400
Db	65937	CGCATATCAATATGAACTGGCTGAAACCCAGAGATGTCAGAGACATGTTTAGTGA	65936
QY	2401	GATCTCTTGGTAAACAATGATGAGCCGGTGAATGTTGAGCCCTGTGTCTTTATTTGCAAGAGCT	2460
Db	65997	GATCTCTTGGTAAACAATGATGAGCCGGTGAATGTTGAGCCCTGTGTCTTTATTTGCAAGAGCT	66056
QY	2461	GTAAGACCTCTGAAAACTTGAACCTAATGATTATGCAATTGAGCTGGGCTTCAGAG	2520
Db	66057	GTAAGACCTCTGAAAACTTGAACCTAATGATTATGCAATTGAGCTGGGCTTCAGAG	66116
QY	2521	AGTAGCGTTGATGAAACTACTGTGTGAAATGTCGGTGCATATGTTAAAGAGGCAAGT	2580
Db	66117	AGTAGCGTTGATGAAACTACTGTGTGAAATGTCGGTGCATATGTTAAAGAGGCAAGT	66176
QY	2581	GTCGAAGATCCCTAGCTGCTGGTGTGGCAATTTGACATGATTCACTGTTCAAGCCAGAGAT	2640
Db	66177	GTCGAAGATCCCTAGCTGCTGGTGTGGCAATTTGACATGATTCACTGTTCAAGCCAGAGAT	66236
QY	2641	TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGAGTTTCTTATGGAATCTGAT	2700
Db	66237	TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGAGTTTCTTATGGAATCTGAT	66296
QY	2701	GTGCGTACCAATAGATGATTAATGATGCAATTTCAATATATGCGATATGTCACAAATA	2760
Db	66297	GTGCGTACCAATAGATGATTAATGATGCAATTTCAATATATGCGATATGTCACAAATA	66356
QY	2761	TGCTTGTTTTGTAGAGCTAAGAACATAGTTCCCATTAATACATGTCCTCCAAAGTTGTACC	2820
Db	66357	TGCTTGTTTTGTAGAGCTAAGAACATAGTTCCCATTAATACATGTCCTCCAAAGTTGTACC	66416
QY	2821	AAGATTAAACAAGTTGCTGAGTAATTTCACTAATTATGCTGCTGTAATTTTTGATCAAA	2880
Db	66417	AAGATTAAACAAGTTGCTGAGTAATTTCACTAATTATGCTGCTGTAATTTTTGATCAAA	66476
QY	2881	CTGTAGACAAATATGTAATTTCACTCTCAACCTTCTGTTTAGAATAACGTAGATTAG	2940
Db	66477	CTGTAGACAAATATGTAATTTCACTCTCAACCTTCTGTTTAGAATAACGTAGATTAG	66536
QY	2941	AGATTGCGCTTAGAGTGGCTTTGTCCAACTTTCTTCTTGAATTTTCTTTGCAATT	3000
Db	66537	AGATTGCGCTTAGAGTGGCTTTGTCCAACTTTCTTCTTGAATTTTCTTTGCAATT	66596
QY	3001	AGGCTCAGTCAGAGCTGACGATTCAGAAAGCATTTCCCAAGATGATGCTAGAGCTGACAG	3060
Db	66597	AGGCTCAGTCAGAGCTGACGATTCAGAAAGCATTTCCCAAGATGATGCTAGAGCTGACAG	66656
QY	3061	GAATATATGATTCAGATGTGAGGAAGATTAAGTCTGTGGCTTTTGCGCCCTGATCACCGAT	3120
Db	66657	GAATATATGATTCAGATGTGAGGAAGATTAAGTCTGTGGCTTTTGCGCCCTGATCACCGAT	66716
QY	3121	AGAAATGTTACAGAGGTGAGGGAATTAATCTACATTCATCAATTTGTGAAAACTGT	3180
Db	66717	AGAAATGTTACAGAGGTGAGGGAATTAATCTACATTCATCAATTTGTGAAAACTGT	66776
QY	3181	TGGAATGATTAATAGTCTGTGTCCTTGTGTGATTCGTATTAATAGTTTGGATGGCC	3240
Db	66777	TGGAATGATTAATAGTCTGTGTCCTTGTGTGATTCGTATTAATAGTTTGGATGGCC	66836
QY	3241	GAATGCTGAAGATTTGGACTGACAGAGCAGCTGAAACTGCGCAGCTGGGTTGCTTTATG	3300
Db	66837	GAATGCTGAAGATTTGGACTGACAGAGCAGCTGAAACTGCGCAGCTGGGTTGCTTTATG	66896
QY	3301	ATTATATACATGTTGAAACTATCTGTTGACAGGTGACAGTCTCAGAGAGATGGAACCGTG	3360
Db	66897	ATTATATACATGTTGAAACTATCTGTTGACAGGTGACAGTCTCAGAGAGATGGAACCGTG	66956
QY	3361	CTCTGTGTGAAGCAACTGTGAGAGATCTGCTGTCTATCTGATTTGGTTCAATCCAGAAA	3420
Db	66957	CTCTGTGTGAAGCAACTGTGAGAGATCTGCTGTCTATCTGATTTGGTTCAATCCAGAAA	67016
QY	3421	ACAATGCTACTGATGTCAGAACCTACACAAAGATAGCAAGTTTTCTGTGTCAAAGTCAG	3480

Db	67017	ACAAATGCTACTGATGTCAAGAACCTTACACAACAGATTCGAAGTTTCTTGTCCTCATCAG	67076
Oy	3481	GGTGGAAAATCACTGAAGGCTGTGTTCTTGSCATCATATAATTAATCAATATGTAGCATGTCT	3540
Db	67077	GGTGGAAAAATCACTGAAGGCTGTGTTCTTGSCATCATATAATTAATCAATATGTAGCATGTCT	67136
Oy	3541	GAGCTTGCGAGATTCTCTTTGTTCTGTAAATTTCTCTCTCTAAGTAGTGTATAAATGA	3600
Db	67137	GAGCTTGCGAGATTCTCTTTGTTGTTGTAAATTTCTCTCTCTAAGTAGTGTATAAATGA	67196
Oy	3601	ACACAAAAAATTAATTAAGCTTCTTGGCACACCCTTTCCCTATCTTAACATATAACATPAAGS	3660
Db	67197	ACACAAAAAATTAATTAAGCTTCTTGGCACACCCTTTCCCTATCTTAACATATAACATPAAGS	67256
Oy	3661	GCTACAA 3667	
Db	67257	GCTACAA 67263	
RESULT 4 AY091075 LOCUS		2637 bp mRNA linear PLIN 18-SEP-2002	
DEFINITION	Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.		
ACCESSION	AY091075		
VERSION	AY091075.1	GI:20259550	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Bukacinski, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	Yamada,K., Banh,J.J., Chan,M.M., Chang,C.H., Dale,E., Deng,J.M., Dengo,J.M., Goldemich,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayaishizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Natusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Arabidopsis Full Length cDNA Clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2637)		
AUTHORS	Yamada,K., Banh,J.J., Chan,M.M., Chang,C.H., Dale,E., Deng,J.M., Dengo,J.M., Goldemich,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayaishizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Natusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-MAR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA		
COMMENT	Riken Genomic Sciences Center (GSC) members carried out the Riken collection and clustering of RAPD cDNAs (RAPL cDNA : 'Riken Arabidopsis Full-length cDNA') : Seki,M., Natusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayaishizaki,Y. and Shinozaki,K.		

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada,K., Banh,J.J., Chan,M.M., Chang,C.H., Chang,E., Dale,E., Deng,J.M., Goldsmith,A.D., Dee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyer,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shim,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)

contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

Location/Qualifiers

```

1. .2637
   /organism="Arabidopsis thaliana"
   /mol_type="mRNA"
   /db_xref="taxon:3702"
   /chromosome="5"
   /clone="RAF109-76-G11 (R19395)"
   /ecotype="Columbia"
   /note="This clone is in a modified pBluescript vector
   (pUC-1) as a BamHI/XhoI insert."
1. .2637
   /gene="At5g42480"
1. .113
   /gene="At5g42480"
114. .2519
   /gene="At5g42480"
   /codon_start=1
   /evidence=experimental
   /product="unknown protein"
   /protein_id="AA13895.1"
   /db_xref="GI:20259551"
   /translation="MRALSHVIGTGLSPFOLCRLPATTKLRBSHSTTTI CSAKMD
RLSDRFRITSDSSSPATRTATTTATVSPSPIDRPRHPIDPFOVGAQTHPLT
DGRVRFPEARVSKRPQGFESDIALISRQLQACETLSPNSRREINELGDEBET
VIGIRPMDKRVGALCVLQEGEGEIVLRVGEALIKERLPSFKDQVVLVNAFLADVS
RDMALDDPDPFITGYEFERBALIKLOBEGSALAPIDRAIDETLEITPRYYIELIG
LPIGDPAARKLNGLSGVRIIWSVGGGASLVGDKREKPMNEATLPMTAAEQVDL
FVATPNSNIPASFEVYVALVAQAFIGKKPHLDDADQFOOLQAKMAEII PMAL
LYTRNNNEIDFELERGLCALLIGKDECKMILGSDSDQYNNPALVEVLENSRD
DNDLDPGLCLTLBTWLAGVFPFRFDTKDKFLGDYYDDPMVLSYERVEVQGSFL
AAAAAARAI GAEHVKA SAMQALQKVPSPRTDNRSAEPKOVSTVESVDEKNNVGD
GEEGVFLAEAVRSENPETNDVAILRAGVSSSVDETVEASVADMLKEASVLLAAGV
AIGLIGLFSOKYRIKSSSFQRKDMVSSSDVATIGSVADDSALPRMDARTAEIIV
VTMOKTSLKLAEPDRHRIEMLPYLDGSMUKIWTDPRAELRQGLAYVDYDILLTSDS
VTVSAQTRALVEATLEESACSLDVLHPENNADIVRIYTRIEVFNKSGMKITGSGV
Lag"
2520. .2637
   /gene="At5g42480"

```

3' UTR

ORIGIN

```

Query Match      52.7%; Score 1931; DB 8; Length 2637;
Best Local Similarity 80.4%; Pred. No. 0;
Matches 2616; Conservative 0; Mismatches 5; Indels 632; Gaps 5;

```

QY	368	GATTTAACCTTAATCTACTCAAAATCAAAATTCCTAATAACCTTAGACGACCAACAGCTC	427
Db	1	GATTTAACCTTAATCTACTCAAAATCAAAATTCCTAATAACCTTAGACGACCAACAGCTC	60
QY	428	TTCAATATGTAAACAGAAACAAAGTTTTTGTAGAGGCTTAAAGACACTCCCTGAGAG	487
Db	61	TTCAATATGTAAACAGAAACAAAGTTTTTGTAGAGGCTTAAAGACACTCCCTGAGAG	120
QY	488	CTCTGAGTCAAGTCGGCATTGGTCTCTCCCATTCCAATTATGCGATTACACCGCGCA	547
Db	121	CTCTGAGTCAAGTCGGCATTGGTCTCTCCCATTCCAATTATGCGATTACACCGCGCA	180
QY	548	CGAAGAGCTTCGACGTAAGCCACAAACCTCTACAATTATCTGTCGCGACGAAATGGG	607
Db	181	CGAAGAGCTTCGACGTAAGCCACAAACCTCTACAATTATCTGTCGCGACGAAATGGG	240
QY	608	CCGACCGCTTCTCTCCGCACTTCATTTCACTCCGATTCCTCTCTCTCTCTCTCGCA	667
Db	241	CCGACCGCTTCTCTCCGCACTTCATTTCACTCCGATTCCTCTCTCTCTCTCTCGCA	300
QY	668	CCGCGACACACCGCCAGCTCTGCTCTCTGTCGACCATTTATGATGTCGCCAAGCC	727
Db	301	CCGCGACACACCGCCAGCTCTGCTCTCTGTCGACCATTTATGATGTCGCCAAGCC	360
QY	728	ACGTCGCCATCCCATTTGATTCTACACAGGATTAAGAGCTCAAAACATTTCTTAACG	787


```
Db 361 |ACGCCATCCCAATGATTTTACAGAGTATAGAGCTCAACACATTTCTTAACCG 420|
Qy 788 |ATGGAATCAGAAAGACATTCGAGCTAGGGTTTCGAAAACCCCGCAATTCGGTTTACGC 847|
Db 421 |ATGGAATCAGAAAGACATTCGAGCTAGGGTTTCGAAAACCCCGCAATTCGGTTTACGC 480|
Qy 848 |ACGACGCTTTATATACCCGAGACAGATTTCTTAAGCTCTTGCGAAACTCTGTCTAATC 907|
Db 481 |ACGACGCTTTATATACCCGAGACAGATTTCTTAAGCTCTTGCGAAAACCTGTCTAATC 540|
Qy 908 |CTCGGCTCAGAAAGAGATACATGAAGGCTCTTCTTAAGATGAAGAAAGCTACATCATCA 967|
Db 541 |CTCGGCTCAGAAAGAGATACATGAAGGCTCTTCTTAAGATGAAGAAAGCTACATCATCA 600|
Qy 968 |CTGATGTTCTTGGGATTAAGGTAATTTGCAATTCGSAATATATAAGTTCTTCGTTTAA 1027|
Db 601 |CTGATGTTCTTGGGAT----- 617|
Qy 1028 |TTTCATGAAATTTGATTAAGAGAACTTTTATCTAGTAAGGTTCTCGGGGCTCTCTGT 1087|
Db 618 |-----AAGTTCTCGGTGCTCTCTGT 638|
Qy 1088 |GTATTCAGAAAGGTTGGAGACTGAGATAGTCTTCGGGTTGGAGGCTCTGCTTAAG 1147|
Db 639 |GTATTCAGAAAGGTTGGAGACTGAGATAGTCTTCGGGTTGGAGGCTCTGCTTAAG 698|
Qy 1148 |GAGAGTTTGCCTAAGTCTTTAAGCAAGATGCTTTTAACTTATGATGAGGCTCTTCTC 1207|
Db 699 |GAGAGTTTGCCTAAGTCTTTAAGCAAGATGCTTTTAACTTATGATGAGGCTCTTCTC 758|
Qy 1208 |GATGCTCGAAGGATGCTATGAGCAATGAGTCACTGATTTTATCTGCTTAAGATTT 1267|
Db 759 |GATGCTCGAAGGATGCTATGAGCAATGAGTCACTGATTTTATCTGCTTAAGATTT 818|
Qy 1268 |GTTGAGGAAGCTTTGAAGCTTTTACAGGTAGTTGACTGCTTTGTTGTTATTTAGACG 1327|
Db 819 |GTTGAGGAAGCTTTGAAGCTTTTAC----- 843|
Qy 1328 |TTGGCTTTATTAAGAACTTTCTTGATTTGATCTTGTATTAAGTCTTGTGTAGAGAAA 1387|
Db 844 |-----AGAGAGAA 851|
Qy 1388 |GAGACAGATGACCTTTCGACCGGATTTACGTGCACAAAATGATGAGACTTTGGAAGATC 1447|
Db 852 |GAGACAGATGACCTTTCGACCGGATTTACGTGCACAAAATGATGAGACTTTGGAAGATC 911|
Qy 1448 |ACTCCGGGTTATGCTTGGAGCTACTTGGCTTACCGCTTGATGATTAAGCTGCGAAA 1507|
Db 912 |ACTCCGGGTTATGCTTGGAGCTACTTGGCTTACCGCTTGATGATTAAGCTGCGAAA 971|
Qy 1508 |AGACTAATATGTTTAAACCGGTGCGGAAATTTTGTGTCGTGTGAGAGAGGTGAGCA 1567|
Db 972 |AGACTAATATGTTTAAACCGGTGCGGAAATTTTGTGTCGTGTGAGAGAGGTGAGCA 1031|
Qy 1568 |TCAAGCTCTTGTGGGGGTTTGACCCGTAAGAAAGTTTATGAATGAAGGCGTTTTCGAATG 1627|
Db 1032 |TCAAGCTCTTGTGGGGGTTTGACCCGTAAGAAAGTTTATGAATGAAGGCGTTTTCGAATG 1091|
Qy 1628 |ACAGCTCTGAGCAAGGTATACAGTTTAGATACCTTTTAAATTTCTTTAGCATGATATA 1687|
Db 1092 |ACAGCTCTGAGCA----- 1104|
Qy 1688 |ACTTAGGTTTCTCATTTTAATGATGTTGTGTAGTAGGATCTTTTGTAGCTACCC 1747|
Db 1105 |-----AGTTGATCTTTTGTAGCTACCC 1128|
Qy 1748 |CAAGCAATATTCAGACAGATCATTTGAAGTTTACGAAGTTGCACTTCTTGCTGCTC 1807|
Db 1129 |CAAGCAATATTCAGACAGATCATTTGAAGTTTACGAAGTTGCACTTCTTGCTGCTC 1188|
Qy 1808 |AAGCTTTATTTGTTAAGAGCACACTTTTACAGATGCTGATTAAGCAATTCAGCAAC 1867|
Db 1189 |AAGCTTTATTTGTTAAGAGCACACTTTTACAGATGCTGATTAAGCAATTCAGCAAC 1248|
Qy 1868 |TTGACAGGCTAAGGTATATGAGCTATGAGAGATTCCTCGGATGTTGATGATACACGAAATA 1927|
Db 1249 |TTGACAGGCTAAGGTATATGAGCTATGAGAGATTCCTCGGATGTTGATGATACACGAAATA 1308|
Qy 1928 |ATTGGAGATTAAGCTTGGTCTTGAAGAAAGGGAATCTGTGTGACATGCTTATAGGCAAAATTG 1987|
Db 1309 |ATTGGAGATTAAGCTTGGTCTTGAAGAAAGGGAATCTGTGTGACATGCTTATAGGCAAAATTG 1368|
Qy 1988 |ATGAATCCGTAATGTGGGCTTAGACAGTGAAGATTCACATATATAGGAATTCAGACTA 2047|
Db 1369 |ATGAATCCGTAATGTGGGCTTAGACAGTGAAGATTCACATATATAGGAATTCAGACTA 1428|
Qy 2048 |TTGTGAGTTTGTTTTGAAGAAATTCAAATCGTAATGAATGATCTCCCTGACTAT 2107|
Db 1429 |TTGTGAGTTTGTTTTGAAGAAATTCAAATCGTAATGAATGATCTCCCTGACTAT 1488|
Qy 2108 |GCAATTTGTTGAAGAACTGTGTGGCAGGGGTTGCTTTCTAGGTTCAAGACACCAAG 2167|
Db 1489 |GCAATTTGTTGAAGAACTGTGTGGCAGGGGTTGCTTTCTAGGTTCAAGACACCAAG 1548|
Qy 2168 |ATAAAAAATTTAACTCGGGGACTAATGATGATGCTATGATGTTTGAATCTTGGAAA 2227|
Db 1549 |ATAAAAAATTTAACTCGGGGACTAATGATGATGCTATGATGTTTGAATCTTGGAAA 1608|
Qy 2228 |GAGTGAAGTATGATCAAGGTTCTCTTTAGCTGCTGTCGACATATAGCAAGATTTGAG 2287|
Db 1609 |GAGTGAAGTATGATCAAGGTTCTCTTTAGCTGCTGTCGACATATAGCAAGATTTGAG 1668|
Qy 2288 |CCGAGCATGTTGAAGCTATGCTATGAGGCACTGCAAGAAATTTTCTTCCCGCTATA 2347|
Db 1669 |CCGAGCATGTTGAAGCTATGCTATGAGGCACTGCAAGAAATTTTCTTCCCGCTATA 1728|
Qy 2348 |CAGATGAAACTCGGCTGAACCCAGAGATGTGCAAGACAGGTTTAAAGTATGATCTCG 2407|
Db 1729 |CAGATGAAACTCGGCTGAACCCAGAGATGTGCAAGACAGGTTTAAAGTATGATCTCG 1788|
Qy 2408 |TTGGTAACTATATAGCCGTGATGATGAGCTGCTGCTTTATTTAGCAAGCTGTAAAGC 2467|
Db 1789 |TTGGTAACTATATAGCCGTGATGATGAGCTGCTGCTTTATTTAGCAAGCTGTAAAGC 1848|
Qy 2468 |CTTCTGAAAATTTTGAATTAATGATTAAGCAATTTGAGCTGGGGTCTCAGAGATAGCC 2527|
Db 1849 |CTTCTGAAAATTTTGAATTAATGATTAAGCAATTTGAGCTGGGGTCTCAGAGATAGCC 1908|
Qy 2528 |TTGATGAACTACTGTTGAAGTCCGTTGCTGATATGTTTAAAGGAGCAAGTGTGAGA 2587|
Db 1909 |TTGATGAACTACTGTTGAAGTCCGTTGCTGATATGTTTAAAGGAGCAAGTGTGAGA 1968|
Qy 2588 |TCTAGCTGCTGCTGCTGCAATTTGACTGATTTCACTGTTCAAGCCAGAAATTTTCTTA 2647|
Db 1969 |TCTAGCTGCTGCTGCTGCAATTTGACTGATTTCACTGTTCAAGCCAGAAATTTTCTTA 2028|
Qy 2648 |AAGCAGCTCATTTTTCAAAGCAAGATATGCTTTCTTATGAAATCTGATGTGCTTA 2707|
Db 2029 |AAGCAGCTCATTTTTCAAAGCAAGATATGCTTTCTTATGAAATCTGATGTGCTTA 2088|
Qy 2708 |CCATAGTATGATTAATATGATGCAATTTTCATATATCTGATTCGCAAAATATGCTGT 2767|
Db 2089 |CCA----- 2091|
Qy 2768 |TTTGTAGCTAAGACATAGTTCCCACTTAATACATGTCACAAAAGTTGTACCAAGATTA 2827|
Db 2092 |----- 2091|
Qy 2828 |ACAAATTGCTGAGTAATTTCACTAATTAATGCTGCTGAATTTTGTGATCAAACTGTAGA 2887|
Db 2092 |----- 2091|
Qy 2888 |CAGAAATGTAATTTCACTCTCAACATTTCTGTTTGAATTAAGATTAAGATTAAGATTCG 2947|
Db 2092 |----- 2091|
```


QY	2948	CTTAGTGTGGCTTGTCCAACTTTTCTTCTGTAATTTTTCTTTGCAATTGAGGCA	3007
Db	2092	-----	
QY	3008	GTCAAGCTGACGATTCAGAAACCTTCCAGAAATGATGCTAGACCTGCAGAAATATA	3067
Db	2100	GTCAAGCTGACGATTCAGAAACCTTCCAGAAATGATGCTAGACCTGCAGAAATATA	2159
QY	3068	GTATCCAAAGTGGCAAGAAATTAACTCTCTGGCTTTTGGGCTGATCACCCGATAGAAATG	3127
Db	2160	GTATCCAAAGTGGCAAGAAATTAACTCTCTGGCTTTTGGGCTGATCACCCGATAGAAATG	2219
QY	3128	TTTACCAAGGTGAGGAAATAATCTACAATTCAATCTGTGTGAAATAAGTTGACAT	3187
Db	2220	TTTACCAAGGTGAGGAAATAATCTACAATTCAATCTGTGTGAAATAAGTTGACAT	2226
QY	3188	GATTATATGTCGTGCTGCTGTTTGTGATCTGTATATTAAGTTTGTGATGGGCAATGCT	3247
Db	2227	-----	AGGTTTGTGATGGGCAATGCT
QY	3248	GAAATTTTGGACTGCACAGAGCACTGAAACTGCGACGTTGGGTTGTTATGATTAAC	3307
Db	2249	GAAATTTTGGACTGCACAGAGCACTGAAACTGCGACGTTGGGTTGTTATGATTAAC	2308
QY	3308	ACTGTTGAAATATCTGTGTGACAGTGTGACAGCTTCACAGATGGAACCCGGTCTGCT	3367
Db	2309	ACTGTTGAAATATCTGTGTGACAGTGTGACAGCTTCACAGATGGAACCCGGTCTGCT	2368
QY	3368	GGAAGCAACTCTGAGGAGAGTCTGCTGTCTATCTGATTTGGTTATCCAGAAAACAATGC	3427
Db	2369	GGAAGCAACTCTGAGGAGAGTCTGCTGTCTATCTGATTTGGTTATCCAGAAAACAATGC	2428
QY	3428	TACTGATGTGGAACCTTACACAAACAATATAGAAATTTCTGGTCCAGTCAAGGTGGA	3487
Db	2429	TACTGATGTGGAACCTTACACAAACAATATAGAAATTTCTGGTCCAGTCAAGGTGGA	2488
QY	3488	AATCACTGAAAGGCTCTGTTCTTGGATCATATATATATATATATAGATGCTGAGCTTG	3547
Db	2489	AATCACTGAAAGGCTCTGTTCTTGGATCATATATATATATATATAGATGCTGAGCTTG	2548
QY	3548	CGAGATTCCTCTTGTCTGTAATTTCTCTCTCTAAGTATGTTTATTAATGAACACAA	3607
Db	2549	CGAGATTCCTCTTGTCTGTAATTTCTCTCTCTAAGTATGTTTATTAATGAACACAA	2608
QY	3608	AAAATTAACGTTTC 3620	
Db	2609	AAAATTAACGTTTC 2621	
RESULT 5	AY221469	2438 bp	linear
LOCUS	AY221469	Arabidopsis thaliana division protein (ARCF) mRNA, complete cde;	PLN 05-AUG-2003
DEFINITION	AY221469	nuclear gene for chloroplast product.	
ACCESSION	AY221469		
VERSION	AY221469.1	GI:33436353	
KEYWORDS			
SOURCE	Arabidopsis thaliana (chale crese)		
ORGANISM	Arabidopsis thaliana (chale crese)		
REFERENCE	1 (bases 1 to 2438)		
AUTHORS	Vitha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Erp,H. and OesterYoung,K.W.		
TITLE	ARCF is a J-Domain Protein and an Evolutionary		
JOURNAL	Plant Cell 15 (8), 1918-1933 (2003)		
PUBMED	12897262		
REFERENCE	2 (bases 1 to 2438)		
AUTHORS	Vitha,S., Koksharova,O., van Erp,H., Froehlich,J.E. and OesterYoung,K.W.		

TITLE Direct Submission
JOURNAL Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
source
1. 2438
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/ecotype="Wassiljevskija"
1. 2438
/gene="ARC6"
/note="synonym: At5g42480"
2. 2407
/feature="ARC6"
/codon_start=1
/product="division protein"
/product_id="AAQ18646.1"
/db_xref="GI:33436354"
/translation="MEALSHVIGIGLSFOLCRLPATTKLRSHNSTSTTCSASKMAD
RLSDPSPSSSSSPATATTTATLTVLSLPSPIDLRPHPIPIDFOVGLAOHFLTL
DGIRAPPEARVSKPPQGFESDDLLISRQLDAACETLSNPRSRREYNELGDLDEGNT
VITIDVPDKYKVALCVLQEGEGTEIVLRVEBALKEKRLPKSFQDVVLVAALFLDVS
RDMALDLPDFITGYEVEBEALKILQEGSSSLAPDLRAQIDETLEIETPRYVELLG
LPIDGVAAKRLNGLSERNILWISVGGSSALVGGTIRKFNMEALETPTAAAEQVLD
FVATPNNIPASFEYEVVALVAQAFICKPHLLDAPDAPFOOLQAKYMANIIPAMF
LYTRNNWEIDPGLERGLCALLIGKIDECRMWGLDSBDQYNNPAI VERVLENSND
DNDLDPGLCLELTLTAGVFERFRIDTKDKFKLGDYDDPMVLSYIERVEVQGSPL
AAATATARIAGEHVKASAMQOLAKVPKSRFTDNRSSAPKQVETPVSDEAGVNNVGS
GEGVFLAEIARVSENEFTDYVALRAGVSSSSYDEETVENSVAEMLEKAVKILLAGCV
AIGLISFSOKYFLKSSSSPFRKDMVSSSPDVAIGSVRADSEALPRMDARTAKENI
VSKWOKTKSLAPGPDRHIREMLPEVLDSGPMKIKWTDRAAEPTAOLGIAYDTLLKLSVNS
VTYSADGTRALVEATLDBESACLSDLVHPENNATIDVRYTTRREYFMSKSMKTKESGV
LAS"
474
/misc_difference
474
/feature="ARC6"
/note="results in Gly to Glu substitution"
/replace="a"
misc_difference 1831
/feature="ARC6"
/note="results in Glu to Asp substitution"
/replace="t"
ORIGIN
Query Match 47.8%; Score 1752.8; DB 8; Length 2438;
Best Local Similarity 79.3%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 2; Indels 632; Gaps 5;
QY 480 CATGGAAGCTCTGAGTACGCTGGCAATGTGCTCTCTCCCAATTCATATGACGATACC 539
DB 1 CATGGAAGCTCTGAGTACGCTGGCAATGTGCTCTCTCCCAATTCATATGACGATACC 60
QY 540 ACCGGCGACGACAAAGCTCCGACGTAGCCACAACACCTCTACAACATATCTGCTCCGAC 599
DB 61 ACCGGCGACGACAAAGCTCCGACGTAGCCACAACACCTCTACAACATATCTGCTCCGAC 120
QY 600 CAATATGAGCCGACCGCTCTCTCTCCGACTTCAATTTGACCTCCGATCTCTCTCTCTC 659
DB 121 CAATATGAGCCGACCGCTCTCTCTCCGACTTCAATTTGACCTCCGATCTCTCTCTCTC 180
QY 660 CTTCGCGACCGCGACCGACCGCGACCTCTCGTCTCTCTGCGACCACTATATGATCTGCC 719
DB 181 CTTCGCGACCGCGACCGACCGCGACCTCTCGTCTCTCTGCGACCACTATATGATCTGCC 240
QY 720 CGAAGCGCAAGTCCCACTCCCAATGATTTCTACACAGGTATATAGAGCTCAAAACACATTT 779
DB 241 CGAAGCGCAAGTCCCACTCCCAATGATTTCTACACAGGTATATAGAGCTCAAAACACATTT 300
QY 780 CTTAACCGATGAATCAGAAAGACATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGG 839
DB 301 CTTAACCGATGAATCAGAAAGACATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGG 360
QY 840 TTTCAGCGACGAGCTTTATTCAGCGCGGAGACAGATTTCTCAAGCTGCTTGGCAAACTCT 899

D 361 TTTACGACGACGCTTTAATCAGCCGAGACAGATTCCTTCAGCTCGCGAACTCT 420
Q 900 GTCTAATCTCGGTCTAGAAAGAGTACAATGAAGTCTTGTGATGAGAAAGCTAC 959
D 421 GTCTAATCTCGGTCTAGAAAGAGTACAATGAAGTCTTGTGATGAGAAAGCTAC 480
Q 960 AGTCATCATGAGTCTCTGGATTAAGTAAATTCATGATTCGGAATATTAAGTTCTT 1019
D 481 AGTCATCATGAGTCTCTGGATTAAGTAAATTCATGATTCGGAATATTAAGTTCTT 505
Q 1020 CGTTTAAATTCATGAAATGGAATTAAGAAAGAACTTTTATCTAGTGAAGTTCCTGGG 1079
D 506 -----AAGTTCTG3GG 518
Q 1080 CTCTCTGTGATTCAGAAAGGTGTGAGACTGAGATGATCTTCGAGGTGTGAGGCTC 1139
D 519 CTCTCTGTGATTCAGAAAGGTGTGAGACTGAGATGATCTTCGAGGTGTGAGGCTC 578
Q 1140 TGTCTAAGAGAGGTGCTTAAGTCTTTAAGCAAGATGTGTTTAAATTAAGGCGCTTG 1199
D 579 TGTCTAAGAGAGGTGCTTAAGTCTTTAAGCAAGATGTGTTTAAATTAAGGCGCTTG 638
Q 1200 CGTTTCTGATGTCTGAGAGATGCTATGGAATGCACTGATTTTATTAATCTGTT 1259
D 639 CGTTTCTGATGTCTGAGAGATGCTATGGAATGCACTGATTTTATTAATCTGTT 698
Q 1260 ATGATTTGTTGAAGAGCTTTGAAGCTTTTACAGATGATTTGACTTGTGTAATTT 1319
D 699 ATGATTTGTTGAAGAGCTTTGAAGCTTTTAC----- 731
Q 1320 GACGAGCGTGGCTTTAATAAGAACTTTCATGATTTGATTTGATGAGTCTGTGT 1379
D 732 ----- 731
Q 1380 AGGAGGAAGAGCAATGACCTTGCAACCGATTTTACGTGCAAAATGATGAGACTTGG 1439
D 732 AGGAGGAAGAGCAATGACCTTGCAACCGATTTTACGTGCAAAATGATGAGACTTGG 791
Q 1440 AAGAGATCACTCCGCGTATGTCTTGAAGCTACTTGGCTTACCGCTGTGATGATTAAG 1499
D 792 AAGAGATCACTCCGCGTATGTCTTGAAGCTACTTGGCTTACCGCTGTGATGATTAAG 851
Q 1500 CTGCGAAAGAATTAATGTTTAAAGCGGTGTGCGAAATATTTTGTGTTGTTGAAGAG 1559
D 852 CTGCGAAAGAATTAATGTTTAAAGCGGTGTGCGAAATATTTTGTGTTGTTGAAGAG 911
Q 1560 GTGAGATCAAGCTCTTGTGGGGGTTTGAACCCGTGAAAGTTTAAAGAAAGCGCTTTT 1619
D 912 GTGAGATCAAGCTCTTGTGGGGGTTTGAACCCGTGAAAGTTTAAAGAAAGCGCTTTT 971
Q 1620 TACGAATGACAGCTGTGAGCAGGATATACAGTTAGATACCTTTTAAATTTCTTAAAGC 1679
D 972 TACGAATGACAGCTGTGAGC----- 992
Q 1680 ATGATATTAATTAAGTTTCTCATTTTAAATGATGTGTGTGAGTGAATCTTTTGT 1739
D 993 -----AGTTGATCTTTTGT 1008
Q 1740 AGCTAATCCCAAGAAATTTTCAAGAGATGATTTGAAGTTTACAAATGATGATCTT 1799
D 1009 AGCTAATCCCAAGAAATTTTCAAGAGATGATTTGAAGTTTACAAATGATGATCTT 1068
Q 1800 TGTGCTCAAGCTTTTATGTTGTAAGAGCAACCTTTTACAGATGCTGATTAAGCAAT 1859
D 1069 TGTGCTCAAGCTTTTATGTTGTAAGAGCAACCTTTTACAGATGCTGATTAAGCAAT 1128
Q 1860 CCAAGCAATTCAGAGGCTTAAGTATGCTATGAGATTCCTGCAATGTTGATGATAC 1919
D 1129 CCAAGCAATTCAGAGGCTTAAGTATGCTATGAGATTCCTGCAATGTTGATGATAC 1188
Q 1920 ACGGAATTAATTTGGAGATAGACTTCGCTTAAGAAAGGGGACTCTGTGCACTGCTTAATAG 1979
D 1189 ACGGAATTAATTTGGAGATAGACTTCGCTTAAGAAAGGGGACTCTGTGCACTGCTTAATAG 1248

Q 1980 CAAAGTTATGAATGCCGTATGTGTTGGCTTGAACAATGAGATTCACAAATATAGAA 2039
D 1249 CAAAGTTATGAATGCCGTATGTGTTGGCTTGAACAATGAGATTCACAAATATAGAA 1308
Q 2040 TCCAGCTATGTGAGAGTTTGTTTTGAAGAAATTAATGCTGATGACATATGATCTCC 2099
D 1309 TCCAGCTATGTGAGAGTTTGTTTTGAAGAAATTAATGCTGATGACATATGATCTCC 1368
Q 2100 TGAATATGCAAAATTTGTGAAACCTGTGTGGAGGGGTGTCTTCTTATGTTTCAAGA 2159
D 1369 TGAATATGCAAAATTTGTGAAACCTGTGTGGAGGGGTGTCTTCTTATGTTTCAAGA 1428
Q 2160 CACCAAGATTAATAATTTTAACTGCGGGAATTAATGATATCTATGTTTGAATTA 2219
D 1429 CACCAAGATTAATAATTTTAACTGCGGGAATTAATGATATCTATGTTTGAATTA 1488
Q 2220 CTTGAAAAGTGAAGTTCAGAGGTTCTCTTTAGCTGTCTGCAATATGCGCAG 2279
D 1489 CTTGAAAAGTGAAGTTCAGAGGTTCTCTTTAGCTGTCTGCAATATGCGCAG 1548
Q 2280 GATTGAGCCGACATGTGAAGCTATGTCTATGCAAGGCACTGCAAAAGTTTCTTC 2339
D 1549 GATTGAGCCGACATGTGAAGCTATGTCTATGCAAGGCACTGCAAAAGTTTCTTC 1608
Q 2340 CCGCTATACAGATAGAAACTCGCTGAAACCAAGATGTGCAAGAGACAGTGTTAAGT 2399
D 1609 CCGCTATACAGATAGAAACTCGCTGAAACCAAGATGTGCAAGAGACAGTGTTAAGT 1668
Q 2400 AGATCTGTGTTGAACAATATGAGCCGTATGTGAGGCTGTGTCTTTATGCAAGAC 2459
D 1669 AGATCTGTGTTGAACAATATGAGCCGTATGTGAGGCTGTGTCTTTATGCAAGAC 1728
Q 2460 TGTAAAGCCCTCTGAATACTTGAATTAATGATTAATGCAATTCAGCTGGGCTCAGA 2519
D 1729 TGTAAAGCCCTCTGAATACTTGAATTAATGATTAATGCAATTCAGCTGGGCTCAGA 1788
Q 2520 GAGTAGGTTGATGAACATACTGTGAATATGTCGTTGCTGATATGTTAAAGAGCAG 2579
D 1789 GAGTAGGTTGATGAACATACTGTGAATATGTCGTTGCTGATATGTTAAAGAGCAG 1848
Q 2580 TGTGAAGATCTTACGCTGTGTGTGCAATGGAATGATTTTCACTGTTACGCAAGAA 2639
D 1849 TGTGAAGATCTTACGCTGTGTGTGCAATGGAATGATTTTCACTGTTACGCAAGAA 1908
Q 2640 TTTTCTTAAAGAGCAATCTTTTCAACGCAAGATATGTTCTTATGGAATCTGA 2699
D 1909 TTTTCTTAAAGAGCAATCTTTTCAACGCAAGATATGTTCTTATGGAATCTGA 1968
Q 2700 TGTGCTAACCATAGATATGATTAATGATGCAATTTTCAATATCTGCAATGCTCAAAAT 2759
D 1969 TGTGCTAACCA----- 1979
Q 2760 ATGCTGTTTGTGAGCTAAGAACATATGTTCCCATTAATATATGTCCTCAAAAGTTGATC 2819
D 1980 ----- 1979
Q 2820 CAAGATTAACAATGTGCTAGTAATTTTCACTAATATATGCTGTGAATTTTGTATCA 2879
D 1980 ----- 1979
Q 2880 ACTGTAGACAGAAATGTAATTTTCACTCTCAACATTTCTGTTTGAATTA 2939
D 1980 ----- 1979
Q 2940 GAGATTGCTTATGCTGTGCTTGTCCAACTTTTCTTCTTGAATTTTCTTTTGCATT 2999
D 1980 ----- 1979
Q 3000 TAGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGGACTGAC 3059
D 1980 TAGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGGACTGAC 2039

QY	3060	GGAATATAGTATCCAACTGGCAGAAGATTAGTCGTGCGCTTTTGCGGCATCACGCCA	3119
Db	2040	AGAAATATAAGTATCCAAGTCGACAGAAAGATTAAAGTCCTGGSCCTTGATCAACCACA	2099
QY	3120	TAGAAAATGTTAACCCAGAGGTGAGGGGAATAAATCTACATTCAATCATATTTGTGAAAACTG	3179
Db	2100	TAGAAAATGTTAACCG-----	2114
QY	3180	TTGGACATGATTATATAGTCTGGTCCTTTGATTTGATTATTAANGTTTTGATGGG	3239
Db	2115	-----AGGTTTTGATGGG	2128
QY	3240	CGAATGCTGAAGATTGTTGACCTGCACAGACAGCTGAAACCTGCGCACTTGGGTTGTTAT	3299
Db	2129	CGAATGCTGAAGATTGTTGACCTGCACAGACAGCTGAAACCTGCGCACTTGGGTTGTTAT	2188
QY	3300	GATTATACACTGTGTGAATACTATCTGTTGACAGSTGTGACAGTCTCAGCAGATGGAAACCCGT	3359
Db	2189	GATTATACACTGTGTGAATACTATCTGTTGACAGSTGTGACAGTCTCAGCAGATGGAAACCCGT	2248
QY	3360	GCTCTGTGTGAAGCACTCTGAGAGAGCTTGTGTCATCTGATTTGTTGTTCAATCCAGAA	3419
Db	2249	GCTCTGTGTGAAGCACTCTGAGAGAGCTTGTGTCATCTGATTTGTTGTTCAATCCAGAA	2308
QY	3420	AACAAATCTATCTGATGTGCAGAACCTTACACAAACAAAGATGGAAGTTTCTGTCCTCAAGTCA	3479
Db	2309	AACAAATCTATCTGATGTGTGCAGAACCTTACACAAACAAAGATGGAAGTTTCTGTCCTCAAGTCA	2368
QY	3480	GGGTGAGAAATACACTGAAGGCTCTGTGTCATCATPATATCTCATATATATGATGACATGTC	3539
Db	2369	GGGTGAGAAATACACTGAAGGCTCTGTGTCATCATPATATCTCATATATGATGACATGTC	2428
QY	3540	TGAAGCTTGCG	3549
Db	2429	TGAGCTTGCG	2438
RESULT 6 AY150490			
LOCUS			
DEFINITION			Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.
ACCESSION			AY150490
VERSION			AY150490.1
KEYWORDS			PLI CDNA.
SOURCE			Arabidopsis thaliana (thale cress)
ORGANISM			Arabidopsis thaliana
REFERENCE			Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyte; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 2436) Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuang,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carminci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,D.P., Southwick,A., Tripp,M.G., Wu,T., Shinzaki,K., Davis,R.W., Ecker,J.R. and Theologis A. Ecker J.R. and Theologis A. Arabidopsis Open Reading Frame (ORF) Clones Unpublished 2 (bases 1 to 2436) Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuang,V.W., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carminci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,D.P., Southwick,A., Tripp,M.G., Wu,T., Shinzaki,K., Davis,R.W., Ecker,J.R. and Theologis A. Direct Submission Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : RIKEN
JOURNAL REFERENCE AUTHORS			
TITLE			
JOURNAL			
COMMENT			

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUN1 (ORF) clones using the RAFL cDNA: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Heuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Tortum, M., Wallender, E.K., Mong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shim, P., Southick, A., Tzip, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC) contributed equally to this work as pIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

location/Qualifiers

1. 2436

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="U19395"

/ecotype="Columbia"

/note="This clone is in PUN1 51."

1. 2436

/gene="At5g42480"

1. 2406

/gene="At5g42480"

/codon_start=1

/evidence="experimental"

/product="unknown protein"

/protein_id="AA12907.1"

/db_xref="GI:23297712"

/translation="MEALSHVGTGLSPFOLCLPAPATYKLRSHNSTSTCSAKMADRLSDSFIDSDSSSFATATTTATLTPSPSIDRPRAPYIPIDFQVGLAQCHPLITDGLRRAPKARVSKPQGFSDALISRQLQAACETLSNPSRRENEGLDDEKATVIDPMDKVPKALCVLQEGEETIVLVEBALKEKLPKSPKQDVLVLAFLDVSRLDAAADDPDPTGYEVEBALKLOEGASLAPDLPAQIDETLETITRYVELLGLGAGGVFAEYAPSENETDYAIVRAGSSSVDETVEVSADMLKEASVYKILAAGVAGLISLPSOKYPLKSSSFORDMVSMSDVATTGSRVADSEALPRMDARPAENIVSKRQKTKSLAFGEEDRIEMLPVLDGRMLKINTDRAEFGQLGYVDYTLILKLSVSVTSADTRALVEATLEESACLSLDVHPENNATDVRYYTTRVEVMSKSMKITEGSLAS"

2407. 2436

/gene="At5g42480"

2433. 2434

/gene="At5g42480"

/note="compared to genomic sequence"

/replace="ctg"

ORIGIN

Query Match 47.6%; Score 1744.6; DB 8; Length 2436;

Best Local Similarity 79.2%; Pred. No. 0;

Matches 2429; Conservative 0; Mismatches 4; Indels 632; Gaps 5;

481 ATGGAAGCTCTGAGTACGATCGGCAATGGTCTCTCCCATTCGAATTATGCGGATTACA 540

1 ATGGAAGCTCTGAGTACGATCGGCAATGGTCTCTCTCCCATTCGAATTATGCGGATTACA 60

541 CCGGCGACGACAAAGCTCCGACGTAGCCACAACAACCTCTACAACTATCTGCTCGGCAGC 600

Db 61 CCGGCGACGACAAAGCTCCGACGTAGCCACAACAACCTCTACAACTATCTGCTCGGCAGC 120

601 AAATGGGCGACGCTCTTCTCTCCGACTTCGAATTCAATTTCACCTCCGATTCTCTCTCTCC 660

Db	121	AAATGGGCGGACCCGCTTCTCTCCGACCTTCAATTTCACCTCCGATTCCTCCTCCCTCC	180
Qy	661	TTTCGCCACCGGCACCAACCAACCGGCACCTCTCGTCTCTGGCCACCATCTATTGATGCTGCC	720
Db	181	TTTCGCCACCGGCACCAACCAACCGGCACCTCTCGTCTCTCCGCCACCATCTATTGATGCTGCC	240
Qy	721	GAAAGCCACGCTCCCATCCCATCCCATTTTCTATCCAGAGTTATAGAGGCTCAACACATTC	780
Db	241	GAAAGCCACGCTCCCATCCCATCCCATTTTCTATCCAGAGTTATAGAGGCTCAACACATTC	300
Qy	781	TTTAAACGATGGAATCAAGAAAGACATTTCCGAGCTAGGGTTTCGAAACCCCGCAATTCGGT	840
Db	301	TTTAAACGATGGAATCAAGAAAGACATTTCCGAGCTAGGGTTTCGAAACCCCGCAATTCGGT	360
Qy	841	TTTCAGCGACGACGCTTTTAATCAGCCGGAGACAGATTTTCAAGCTGCTTCGGAAACTCTG	900
Db	361	TTTCAGCGACGACGCTTTTAATCAGCCGGAGACAGATTTTCAAGCTGCTTCGGAAACTCTG	420
Qy	901	TCTTAATCCTCGGCTCTTGAAGAGAGATCAATGAAGTCTTCTTGATGATGAAGAGATCA	960
Db	421	TCTTAATCCTCGGCTCTTGAAGAGAGATCAATGAAGTCTTCTTGATGATGAAGAGATCA	480
Qy	961	GTCATCATCTGATGTTCTCTGGGATTAAGGTAATTGCAATTCGGAAATAAAGTTCTTTC	1020
Db	481	GTCATCATCTGATGTTCTCTGGGATTAAGGTAATTGCAATTCGGAAATAAAGTTCTTTC	504
Qy	1021	GTTTAAATTTCAATGAATTGGAATGAAGAAAGAACTTTATCTAGTAAAGTTCTCGGGGC	1080
Db	505	-----AAGTTCTCGGGGC 518	
Qy	1081	TCTCTGTGTATTGCAAGAGGTGTGTGACATGATTAAGTTCTCGGGTTGGTAGGCTCT	1140
Db	519	TCTCTGTGTATTGCAAGAGGTGTGTGACATGATTAAGTTCTCGGGTTGGTAGGCTCT	578
Qy	1141	GCTTAAGAGAGAGGTTGCCTAAGTCGTTTAAAGCAAGATGAGTTTAAAGTATGAGCGCTGC	1200
Db	579	GCTTAAGAGAGAGGTTGCCTAAGTCGTTTAAAGCAAGATGAGTTTAAAGTATGAGCGCTGC	638
Qy	1201	GTTTCTCGATGTCTCGAGGAGATGCTATGGCAATTGATCAACCTGATTTTATTACTGTTA	1266
Db	639	GTTTCTCGATGTCTCGAGGAGATGCTATGGCAATTGATCAACCTGATTTTATTACTGTTA	698
Qy	1261	TGAGTTTGTGAGAAAGCTTTGAAGCTTTTACAGGTAAGTGAATGCTTGTGTAATTG	1320
Db	699	TGAGTTTGTGAGAAAGCTTTGAAGCTTTTACAGGTAAGTGAATGCTTGTGTAATTG	730
Qy	1321	ACGAGCGTTGGCTTATAAAGACTTCTTGAATTGTAATCTTTGTAATTGAAGCTTGTGTA	1380
Db	731	-----A 731	
Qy	1381	GGAGAGAGAGCAAGTAGAGCTTGCACCGAATTTACGTGCACAAATGATGAGACTTTGGA	1440
Db	732	GGAGAGAGAGCAAGTAGAGCTTGCACCGAATTTACGTGCACAAATGATGAGACTTTGGA	791
Qy	1441	AGAGATCACTCCGCGTTATGTCTTGAGAGTACTTGGCTTACCGCTTGGTGAATTAACG	1500
Db	792	AGAGATCACTCCGCGTTATGTCTTGAGAGTACTTGGCTTACCGCTTGGTGAATTAACG	851
Qy	1501	TGCGAAAGACTAATNGTTTAAAGGGGTGTGCGAAATTTTGTGCTGTGTAAGAGAGG	1566
Db	852	TGCGAAAGACTAATNGTTTAAAGGGGTGTGCGAAATTTTGTGCTGTGTAAGAGAGG	911
Qy	1561	TGAGACATCAGCTCTTGTGTGGGGTTTGAACCGGTGAGAGTTTATGATGAGCGCTTTT	1620
Db	912	TGAGACATCAGCTCTTGTGTGGGGTTTGAACCGGTGAGAGTTTATGATGAGCGCTTTT	971
Qy	1621	ACGAATGACAGCTGTGACAGGTAATACGATTAGATCCTTTTAAATTTCTTACGA	1680
Db	972	ACGAATGACAGCTGTGAGC----- 991	
Qy	1681	TGATATACTTAAAGTTTCTCATTTTAAATGATGTGTGTGATGATGATCTTTTGTGA	1740
Db	992	-----AGTTGATCTTTTGTGA 1008	

Oy	1741	GCATCCCCAAGCAATATTCACGACGAGTCAATTGAAATTACGAAGTTGCATCTGCTTT	1800
Ds	1009	GCTACCCCAAGCACAATATTCACGAGAGTCAATTGAAATTACGAAGTTGCACTTGCTCTT	1068
Oy	1801	GTGCTCAAGCTTTTATTTGGTAAGAACCCACCTTTTACAGGATGCTGAATTAAGCAATTC	1860
Ds	1069	GTGCTCAAGCTTTTATTTGGTAAGAACCCACCTTTTACAGGATGCTGAATTAAGCAATTC	1128
Oy	1861	CAGCAACTTCACGAGGCTAAGTAATGCTATATGAGATTCCTGCAATGTTGATGATACA	1920
Ds	1129	CAGCAACTTCACGAGGCTAAGTAATGCTATATGAGATTCCTGCAATGTTGATGATACA	1188
Oy	1921	CGAATTAATTTGGAGATAGACTTCGCTCTAGAAAGGGAACTCTGTGCACTGCTTATAGGC	1980
Ds	1189	CGAATTAATTTGGAGATAGACTTCGCTCTAGAAAGGGAACTCTGTGCACTGCTTATAGGC	1248
Oy	1981	AAACTTATATGAATGCCGATATGCTGTGGGCTTGAACATGATAGATTCACAATATATAGAAAT	2040
Ds	1249	AAACTTATATGAATGCCGATATGCTGTGGGCTTGAACATGATAGATTCACAATATATAGAAAT	1308
Oy	2041	CCAGCTATTTGAGAGTTGTTTGGAGAAATTCATAAATCGTATGACATATATCTCCCT	2100
Ds	1309	CCAGCTATTTGAGAGTTGTTTGGAGAAATTCATAAATCGTATGACATATATCTCCCT	1368
Oy	2101	GGACTATGCAAAATTTGTGGAAACCTGTGTGGCAGGGGTGTCTTTCTTAGTTCAAGAC	2160
Ds	1369	GGACTATGCAAAATTTGTGGAAACCTGTGTGGCAGGGGTGTCTTTCTTAGTTCAAGAC	1428
Oy	2161	ACCAAAATATAAAAATTTAAATCTCGGGGACATACATATGATCTCTATAGTTTGAATTAC	2220
Ds	1429	ACCAAAATATAAAAATTTAAATCTCGGGGACATACATATGATCTCTATAGTTTGAATTAC	1488
Oy	2221	TTGGAAGAGTGAAGATGCTTACAGGGTCTCTTTAGCTGCTGTCGACATATGGCAGG	2280
Ds	1489	TTGGAAGAGTGAAGATGCTTACAGGGTCTCTTTAGCTGCTGTCGACATATGGCAGG	1548
Oy	2281	ATTGAGACCCAGCATGTGAAAGCTATGCTATATGACGCACTGCAGAAAGTTTTCTCTCC	2340
Ds	1549	ATTGAGACCCAGCATGTGAAAGCTATGCTATATGACGCACTGCAGAAAGTTTTCTCTCC	1608
Oy	2341	CGCTATACATATGAAATCTGGCTGAAACCCAGGATGTGCAAGACAGTGTTAAGTGA	2400
Ds	1609	CGCTATACATATGAAATCTGGCTGAAACCCAGGATGTGCAAGACAGTGTTAAGTGA	1668
Oy	2401	GATCTCTTGGTAAACAATGTAGGCGCTGATATGATGAGCCTGTGTCTTATTTGCAAGACT	2460
Ds	1669	GATCTCTTGGTAAACAATGTAGGCGCGTATGATGAGCCTGTGTCTTATTTGCAAGACT	1728
Oy	2461	GTAAGACCTCTCTGAAAATTTGAAACTTAATGATTTATGCAATTCGAGCTGGGCTCTCAG	2520
Ds	1729	GTAAGACCTCTCTGAAAATTTGAAACTTAATGATTTATGCAATTCGAGCTGGGCTCTCAG	1788
Oy	2521	AGTAGCCTTGAATGAATACTACGTGTGAAATGTCCTCGTGTGCTATATGTTAAAGGAGCAGT	2580
Ds	1789	AGTAGCCTTGAATGAATACTACGTGTGAAATGTCCTCGTGTGCTATATGTTAAAGGAGCAGT	1848
Oy	2581	GTGAAGATTCCTAGCTGCTGCTGTGGCAATTTGACATGTTCACTGTTCAAGCAGAAATAT	2640
Ds	1849	GTGAAGATTCCTAGCTGCTGCTGTGGCAATTTGACATGTTCACTGTTCAAGCAGAAATAT	1908
Oy	2641	TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATATGTTTCTTCTATGAAATCTGAT	2700
Ds	1909	TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATATGTTTCTTCTATGAAATCTGAT	1968
Oy	2701	GTGCTTACCATATGATATGATTTAAATGATATGCAATTTTCAATATCTGATTCCTCAAAATA	2760
Ds	1969	GTGCTTACCATATGATATGATTTAAATGATATGCAATTTTCAATATCTGATTCCTCAAAATA	1978
Oy	2761	TGCTTGTTTTGTAGCTATAGAACATATGTTCCACTTAATATCATGTCCCAAAAGTTGTACC	2820
Ds	1979	TGCTTGTTTTGTAGCTATAGAACATATGTTCCACTTAATATCATGTCCCAAAAGTTGTACC	1978

[illegible]

TITLE
 JOURNAL
 COMMENT
 Direct Submission
 Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 On Jan 21, 2004 this sequence version replaced gi:39142429.
 Genes were predicted from the integrated results of the following:
 GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), POENISH
 (http://www.softberry.com/), GeneMark.hmm
 (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
 (http://www.tigr.org/cdb/glimmerm/1mr.form.html), RiceHMM
 (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
 (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4
 (http://glodion.cse.psu.edu/html/docs/sim4.html), gap2
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI NonRedundant Protein
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
 sequence database at RGP or DDBJ. Protein homologies of the coding
 regions were searched against NCBI NonRedundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTN with the corresponding DDBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as a
 probable 'hypothetical' protein and is included as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from 5' to 3' of the PAC clone.
 This sequence of P0575F10.1 clone has an overlap with OJ1020.C02
 (DDBJ: AP004076) clone at 5' end and with P0482P12 (DDBJ: AP005311)
 (DDBJ: AP004076) clone at 3' end. Detailed information on overlap and assembly
 quality together with annotation of this entry is available at
 http://rgp.dna.affrc.go.jp/GenomeSeq.html.
 location/Osaffr1
 1.150462
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="P0575F10"
 /complement(4987..8112)
 /gene="P0575F10.1"
 /complement(join(4987..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..8112))
 /gene="P0575F10.1"
 /note="supported by full-length cDNA(s): AK106553"
 /complement(join(5003..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..8112))
 /gene="P0575F10.1"
 /note="supported by full-length cDNA(s): AK073059"
 /complement(join(5315..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..7876))
 /gene="P0575F10.1"
 /note="contains EST(s): AU070160(E61534)
 contains full-length cDNA(s): AK106553,AK073059"
 /codon_start=1
 /product="putative eukaryotic translation initiation
 factor 3, subunit 1 alpha, 35kDa"
 /protein_id="BAD07941.1"
 /db_xref="GI:41053010"
 /translation="MEDWSEDFQPAVPAPAKAPLKSVMADVEDBDDVKSWESEEE
 EKRPVPEVKTAPEKSGAKKGGKQOASVSSVSEVQEDLADPELTKLQROALVEAD
 FKSTTLRFGKDGSEKSLDFPIKSSDPAFVEILANKEALPYEKSPHYWGLKNWR
 LSNASLKGADAKDQISSIALINAEKTKAEEMAAAGKKKQAKKQHLIENKDDFIG

gene KGNFDDPDEYDEM"
 /complement (10186, .13754)
 /gene="P0575F10.2"
 /complement (join(<10186, .10479,10759, .10892,11127, .11991, 12089, .12349,12515, .12712,12843, .13073,13455, .>13754))
 /gene="P0575F10.2"
 /note="start and end point are not identified"
 /complement (join(10186, .10479,10759, .10892,11127, .11991, 12089, .12349,12515, .12712,12843, .13073,13455, .13754))
 /gene="P0575F10.2"
 /note="contains EST(s) : AU058443 (E51136), AU183658 (E51136)"
 /codon_start=1
 /product="placitid division protein precursor"
 /protein_id="BAD07942.1"
 /db_xref="GI:41053011"
 /translation="MEGFHNILARPNAGAPAFSLPRPRPRRRPPHPSACRAAS
 WABRLPADPHILTPRAPSDDPSRAPAAPASAPVPLPDDAERSLQVDPYKUL
 GAERHFLGDTIRAFERIRAKPPQYGTSTDLVRRQMQLADHTLNQSRQYDA
 LSEHREALTMIDAMDREAGALAVLTGQLDRPKPKQDVITAMALAYDLR
 DAMAASPDIYGCCEVERALKIQEDGASNLADLSQIDTELETTPRCVLESL
 PITEHHKKEOEGLQCARNILMSYRGGIAVGGGFSREAFMNEAFLMRSIEQMDFF
 SKTPNSIPPEWFELIYNALAVAOAISKRPOFIMMADLFEOLQKFNISHYDNE
 MDALERAFCSLIVGVSKRMWIGDNESPPYDPILEPIITVNSSISENDILPGL
 CKILETWLIEVPRSRDTRGMQRLDIYDDEPVASYLERMEGGSALAAAAIYK
 LGQATVATGIVSNALQAFKVFPLIEQDRSMENTKQPGGILSEFQENAPAD
 SRNALKTIISAGALFALAVIAGAKYLRKQLSIRSHGSVAANVSDDTDPALDE
 DPVHPRMDAKLADIYRKQSIKSKALGPEHSVASIQEVLQGNMLKVTMDRAEIR
 HGMWEYTLSDVTITSLIDGRRAVEATIDBAQGLTVTPEPRNDSYDTKYTRY
 EMAPSKCGMKITGGAVLKS"
 complement (16401, .17341)
 /gene="P0575F10.3"
 /complement (join(<16401, .16739,16841, .>17341))
 /gene="P0575F10.3"
 /note="start and end point are not identified"
 /complement (join(16401, .16739,16841, .17341))
 /gene="P0575F10.3"
 /note="similar to Arabidopsis thaliana chromosome 3, At3g61870"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD07943.1"
 /db_xref="GI:41053012"
 /translation="WMAPIAAVAHAIITCSASKNSPSARQQOQTTTATRGSPA
 LPSILRTTAAATAALALAPDPAALAGGEGILGGSVALHLVWGLFAATLWAG
 YLGQMRERVTIODEINELKKQLPAAALAPPAVAADSSSSSPAPAPSPVVI
 KIDELTERKKLIGKSPFRDHFNAGSTLLGIVTESGVALNTMFRGKLPFGHLEA
 GAATVMAAAALVPMOKGNERYASLHAIANINVLPTWQIPGTELVGKVFERT
 TWP"
 complement (18499, .18942)
 /gene="P0575F10.4"
 /complement (<18499, .>18942)
 /gene="P0575F10.4"
 /note="start and end point are not identified"
 /complement (18499, .18942)
 /gene="P0575F10.4"
 /note="contains EST(s) : AU184851 (R10374), AU184850 (R10374)"
 /codon_start=1
 /product="putative BF-hand Ca2+-binding protein CCD1"
 /protein_id="BAD07944.1"
 /db_xref="GI:41053013"
 /translation="MNHQIIRSTTAAEQEASAGGGGSEYEDLMFVAGRLGAEGL
 LSEIRAGFLADPARGAITAEISLRSAAVLIGGGGGEYVEAAANVAREGDODD
 GAISEAFVYLVTLSPGIMDAGMLEAIADELASLPPPPA"
 21937, .23552
 /gene="P0575F10.5"
 /join(<21937, .22225,22326, .22648,22736, .23049, 23393, .>23552)
 /gene="P0575F10.5"
 /note="start and end point are not identified"
 /join(21937, .22225,22326, .22648,22736, .23049,23393, .23552)
 /gene="P0575F10.5"
 /note="similar to Arabidopsis thaliana chromosome 5, At5g24330"

Query Match 8.7%; Score 317.6; DB 8; Length 150462;
 Best Local Similarity 49.3%; Pred. No. 3.3e-64;
 Matches 1459; Conservative 0; Mismatches 1299; Indels 202; Gaps 16;
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD07945.1"
 /db_xref="GI:41053014"
 /translation="MGAPPLRRRTARPARATREAGSGGDDDDVDRCEAGSGESA
 AELLDCDGRGHICLRPLIPVPAQMFCSQSPSHSKSKSAKXKQPLVVO
 TKIVDFKIDRGPAALAAESSBEGKRRKRGIRLVSKKRLPLFPSPDDPAR
 LRQMASLATLVTGAVFSENLTYVQMAPRANRALESQVQVLPKEDVETLNCK
 RMAARGPEPLIVVDVEGFVEADRFIDLTITREYVDVLTFRREDDDSMT
 LLSAATPSESLVCPDKRSNIARFINGINHTFDGRKKQIKCVRFVGGECRVLLVA
 NRDISGERLYDYNGSEHYPTHYV"
 complement (join(24479, .24817,24913, .25266,25398, .25633, 25720, .25832,27054, .27185,27995, .28148))
 /gene="P0575F10.6-1"
 /complement (join(24479, .24817,24913, .25266,25398, .25633, 25720, .25832,27054, .27185,27995, .28148))
 /gene="P0575F10.6-1"
 /note="supported by full-length cDNA(s) : AK121372"
 /complement (join(24491, .24817,24913, .25266,25398, .25633, 25720, .25832,27054, .27185,27995, .28147))
 /gene="P0575F10.6-1"
 /note="supported by full-length cDNA(s) : AK062255"
 /complement (24563, .28147)
 /gene="P0575F10.6-2"
 /complement (24563, .28147)
 /gene="P0575F10.6-2"
 /note="contains full-length cDNA(s) : AK065148
 non-coding transcript
 probably inactive due to including stop codon(s) in CDS"
 complement (join(24787, .24817,24913, .25266,25398, .25633, 25720, .25832,27054, .27159))
 /gene="P0575F10.6-1"
 /note="contains EST(s) : AU031017 (E60571), AU093466 (E60571)
 contains full-length cDNA(s) : AK121372, AK062285"
 /codon_start=1
 /product="putative arginine/serine-rich splicing factor"
 QY 744 TGATTTCTACCAAGTATTTAGAGACTCAACACATTTCTTAACCGATCGAATCGAAGAGC 803
 Db 13086 TGGCTTTGCTCAGGTTCTAGGGGACAGGCCATTTCTTGGGAATGCGATCGAAGAGC 13027
 QY 804 ATTGGAAGTAGGGTTTCGAACCGCGCAATTCGGTTGAGGACGACGCTTTAATCAG 863
 Db 13026 GTTCGAGGACGGAATGCCAGCACCGCAGTATGGCTACAGACGGAATGCTTTGTGG 12967
 QY 864 CCGGAGACAGATTCCTCAAGCTCTTGCMAACTCTGTCTAATCTCGGTCTAGAAAGA 923
 Db 12966 TCGTCGACAAATGCTCGAATGCCATGACATCTCATGAAACGAACTCCGCACTCA 12907
 QY 924 GTTCATGAAGGTCCTTGTATGATGAAGAGCTACAGTATCATCGATGTTCTTGGGA 983
 Db 12906 GTATGATCGGCCCTTTCGAGAACCGTGAAGAGCTTCACATGGAATTTCTGGGA 12847
 QY 984 TAAGTAATTTGATTTCCGAAATATTAAGTTTCT-----TCGT 1022
 Db 12846 CAAGGTGAGGCTTTAGGGGACCAAGGATGTTTGAATCTTTCGAATACACTAAGT 12787
 QY 1023 TTTAATTTCAATGATTTGATTAAGAAAGAACTTTATCTAGTGAAGTTCTCGGGCTC 1082
 Db 12786 GTATTAATCTCTCTGCTTATTAATATATGCGGTCTCTTGGAAGCAGGTTCCAGGTGTG 12727
 QY 1083 TCTGTGATTTGCAAGAGGTGTGAGACATGAGATAGTTCTTGGGTTGGTGGAGCTCTGC 1142
 Db 12726 TGTGTGCGCTTCAGAGGCTCGGAGGACCTTCTGCTGTGTAATCTGAGAAACAGTTGC 12667
 QY 1143 TTAAGGAGGTTGCGTAAGTGTTTAAGCAAGATGATGTTTATAGTATGCGCTTGCGT 1202
 Db 12666 TCTGATTCGCGCACCAAGCGCTTCAAGCAGACGATGTTGCTAGCATGCTTGCTT 12607

QY 1203 TTCTGATGTCGAGGAGTCTGATGCAATTGATCCACTGATTTTATTTACTGTTATG 1262
 DB 12606 ATGTGATCTATCAAGGGATGCTATGGCAGAACCCCTCCAGATGTAATGGCTGCGC 12547
 QY 1263 AGTTTGTGAGGAAGCTTTGAAGCTTTTACAG-----TAG 12398
 DB 12546 AGGTGCTGAGAGGGCTCTCAAGCTCTTGACAGTGAACCTGTTTGTCTTTTACTCCCTAC 12487
 QY 1239 TTTGACTGCTTGGATTTGAGAGCGTTGGCTTTATAGAACTTTCTTGATTTG--- 1355
 DB 12486 TTTTTCATTTGGGCTTATATTTGAATAGATTAAGATTAATTAACCTGTGAAGATTGGTTT 12427
 QY 1356 -----ATACCTTGTATTGAGCTTG----- 1376
 DB 12426 GGGACACATGGGCTATATTTTATTAAGAAATGTAATTTCTTGAAGCAATATATA 12367
 QY 1377 -----TGAAGAGAGAGAGCAAGTACCTTTGACCGGATTTACGTGACAAA 1424
 DB 12366 GGAACCTTCATCTGTAGGAAGATGAGACAACTCGCACTGATCTGCTTTCACAGA 12307
 QY 1425 TTGATGACTTTGGAAGAGATCACTCGCGCTTATGTTCTTGGACTACTTGGCTTACCGC 1484
 DB 12306 TTGATGAACCTCTGAGAGATTTACCTCGCTGTATTTGAGCTTCTCTCCCTTCTTA 12247
 QY 1485 TTGGTGAATGATTAAGCTGCGAAAAGACTAATGTTTAAAGCGGTGCGGAAATATTTGT 1544
 DB 12246 TTGACACAGAGCATCATTAAGAGGCGCCAGAGAGGCTTCAAGGTGCGAGAAATTTTGT 12187
 QY 1545 GGTCTGTTGAGAGAGTGAGACATCACTCTTGTGGGGGTTTGAACCGTGAAGATTTA 1604
 DB 12186 GGAGCGTTGGCAGAGAGGTATGCTACCGTTGAGAGGATTTCTCGTGAACCTTCA 12127
 QY 1605 TGAATGAGCGCTTTTACGAATGACAGTCTGAGCAAGT-----ATACGTTTGA 1655
 DB 12126 TGAACGAGCGCTTTTGAAGATGACATCAATTAAGACAGGTTCGATGTTGTTTAT 12067
 QY 1656 ATACCTTTTATTTCTTGAAGATATPACTTTAGTTTCAATTTAATGATGT 1745
 DB 12066 GAGCAAGCTGCTATTTATTTGAGAGACATCTTTGAAAACATGTTAACTCCCTTC 12007
 QY 1716 TGT-----GTGGTAGTGTGATCTTTTGTAGCTACCCGAGCAATATTCAGCAGAGTCA 1770
 DB 12006 TGTTAACCTCCGAGATGATTTCTTTCAAAAACCGAATACATTTCCCTGAATGG 11947
 QY 1771 TTTGAAGTTTACGAAGTTGCACTGCTCTTGTGCTCAAGCTTTTATTTGTAAGAGCCA 1830
 DB 11946 TTTGAAATTTTACATGATGACATCTGCAATGCTCAGCAATTAATTAAGAAAGGCCA 11887
 QY 1831 CACCTTTTACGATGCGATGCAATTCAGCAACTTCAGAGGCTAAGGTATGGCT 1850
 DB 11886 CAATTCATCATGATGGGAGATCTTTTGAACAACCTCCAGAAATTCACATAGGTTCT 11827
 QY 1891 ATGAGAGATTCCTCGATGTTGTATGATATACAGAAATATTTGGAGATAGACTCGCTTA 1950
 DB 11826 CATTAATGCTTAATG-----ATATGAGATGACCTTGCAATGG 11791
 QY 1951 GAAAGGGGACTCTGTGCACTGCTTATTAAGCAAAAGTTGATGATGCCGATATGTGGCC 2010
 DB 11790 GAAAGGGGACTCTGTGCACTGCTTATTAAGCAAAAGTTGATGATGCCGATATGTGGCC 11731
 QY 2011 TTGACAGTGAAGTCAATATAGAAATTCAGCTATTTGAGGTTTGTGTAAGAT 2070
 DB 11730 ATTGATATATGATTTTCAACATACAGAAATTCAGAAATTTCTAGAGTTTATTTGACCAAC 11671
 QY 2071 TC---AATCGTGAATGCAATGATGATCTCCCTGCACTATGCAAAATTTGGAACCTGG 2127
 DB 11670 TCTAGCATCACTGTAAGAAATGATCTTCTCCAGGGGCTGCAAGCTTTTGAAGCTTGG 11611
 QY 2128 TTGGCAGAGGGTGTCTTCTTCAAGTTCAAGAGACCAAAATTAATAAACTGGGG 2187
 DB 11610 CTATATCTTTGAGGTTTCTTCAAGAGATGATCTCGGGGCAATGCAAGTTTCAAGCTTGG 11551
 QY 2188 GACTACTATGATGATCGATGCTTTTGAAGTTACTTGAAGAAAGTGAAGTACAGGGT 2247

DB 11550 GATTACTACATGATATCCAGAAAGTTTAAAGCTACTAAGAAAGATGAGGGGTGTGCT 11491
 QY 2248 TCTCCTTTAGCTGTGCTGCACTATGAGCAAGATTGGAGC-----C 2289
 DB 11490 TCTCATTTGGCTGTGCTGCTGCTATTTGCAAACTTGGTGTCTAAGCTACAGCTGACTT 11431
 QY 2290 GAGCATGGAAGCTAGTGTGCTATGCAAGCACTGCAAGAAAGTTTCTCCCGCTATACA 2349
 DB 11430 GGTACTGTGAATCAATGCTATTCAGAGCTTCAACAGAGTTTTCATTTGATAGAACAG 11371
 QY 2350 GATGAAACTCGGCTGAACCCAGAGATGTCAGAGACAAGTATTAGTATGATCTGTT 2409
 DB 11370 TTAGACAGGTGACGATGAGAAATATCAAGATGCGCCCTGGGGATATCTGAAATTTT 11311
 QY 2410 GGTACAAATGATGAGCGCTGATGATGAGCTGG-----TGTCTTATTT 2451
 DB 11310 GACCAGAAATGACCTGCTCATGATTCAGAAATGCCCTTGAAGATTTATCTGTCT 11251
 QY 2452 GCAGAGCTGTAAGACCTCTGAAAACCTTGAACCTATATGATTAATGCAATTGAGCTGG 2511
 DB 11250 GGCCACTGTTTGCATCTGTTGGCAGTAAATTTGGGCCAAATATTTGCTCGTAAGAGGCC 11191
 QY 2512 GTCTCAGAGATGACGTTGATGAACCTACTGTTGAATGTCCTGCTGATATGTAAG 2571
 DB 11190 CTTTCTGCTATTTAGAGATGAGATGATCTGTGCAGTGTATATGATGATGATCTACT 11131
 QY 2572 GAGCAAGTGTGAAGATCTTACGCTGCTGTGTGCAATTTGACATGATTTCACTGTACG 2631
 DB 11130 GATGTGATGAATAGCTTTTCATTT--TGTAAATGATGCAACCACTGGAATTTACTG 11073
 QY 2632 CAGAGATTTTCTTAAAGAGCTCATCTTTCAAGCAAGATGATGTTCTTCTATG 2691
 DB 11072 CTACGTTTACTGGGTATCTTGTGGCTGATTAAGTATGTTCTTGAATCTTATAGCT 11013
 QY 2692 GAATCTGATGCTGCTACCATAGTATGATTAATATGATGCAATTTTCATATATCTCATG 2751
 DB 11012 TTGATTAATGTTTAAACAAGATATCATCAAAATTTTGTCTCATATACATAG 10953
 QY 2752 CTCAAAATATGCTGTTTGTGTAGCTTAAGACATATGTTCCACTTAATATACATGCCAA 2811
 DB 10952 CTACTATTAATCTGTATATGATATCATAGAGATTAATTTGATGATGATGATGTAAG 10893
 QY 2812 AGTTGTAACAAAGTAAAGTGTGCTGATGAATTTACTAATTATGCTGTTG----A 2866
 DB 10892 ATCTGCACTAGATGAAGATTCAGTACATATTTCTAAGATGAGCCAGAGTGCAGAG 10833
 QY 2867 ATTTTGTATCAAACTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGAA 2926
 DB 10832 ATATTTGTGCAAGTGCAGATATCAAAATCTAAGGCTTGGACCAAGAACTTGGGTTG 10773
 QY 2927 TAACTGATGATTAAGATGCTTATGATGCTTGTGCAACTTTTCTTTC--CTTGAATTT 2985
 DB 10772 CATCATGCAAGAGGTATGTTTAAACTTTGACTTTTCACTGCTAGTGGCCACTG 10713
 QY 2986 TTTTCTTTTGAATTAAGTCAAGTCAAGAGCTGCAAGTTCAGAACACTTCCAGAAATGGA 3045
 DB 10712 CTCTTACCTTATCTTACGCTTCTATAGAGATTAATTCAGATTAATCTCTTATGGA 10653
 QY 3046 TGTAGACATGCAAGATATATGATTCAGATGTCAGAAAGTATTAAGTCTGCTGCTTTGG 3105
 DB 10652 TATATTTGTGACGCAATGATCTTTTACAGTTTGTATCAATATATCTTTCT----- 10602
 QY 3106 GCTGATACCGCATGAATATGTTTACAGAGAGTGAAGAAATTAATCTACATTCATCA 3165
 DB 10601 ---TGTGATCTCAACACATGCTTCAATTTGCACTGAACAAATTTTCAAGAACTATTTGC 10545
 QY 3166 TTGTGTAAAACGTGTGACATGATTAATATGCTGAGCTTGTGTTGAT---CTGTATTT 3222
 DB 10544 CACTGATGATATCTCCGTATATGATGCGGTCAAAATCTGATGATTTTCAATGTTCT 10485
 QY 3223 TATAGTTTGTGATGAGCGAATGCTGAAGTATTTGATGCTGACAGAGCACTGAAACTGCGC 3282

Db 10484 TATAGTTCCTTGATGCAACATGCTAAAGGTGTGSACTGACCGAGCGAGATTTAGC 10425
Qy 3283 AGCTGGGTTGGTTATGATATATACACTGTGTGAATACTATCTGTGACAGTGTGACAGTCT 3342
Db 10424 GTATGGGTGGTTGTGGAGATATACATATCCGATGTGACGATTTGATGATCATCTATCT 10365
Qy 3343 CAGACAGTGAACCCGTGCTGTGTGTGAGAGCACTGTGAGAGTCTGTGCTTGTCTTATCTG 3402
Db 10364 CCCATGATGTGTGCGACGACGCTGTGAGGCTCAATGATGAGAGCGCCACCTTACTG 10305
Qy 3403 ATTGTGTTCAATCCGAAAACAATGCTATGTCATGACGAAGCTACACAACAGATACGAG 3462
Db 10304 ATGTACTAGCCCAAGAACATATATCATATGACACAAAATACATACCCGGTATGGA 10245
Qy 3463 TTTTCTGTCCAG---TCAGGTGGAATAATCACTGAAGCTGTCTTCTTCATCATATAT 3519
Db 10244 TGGCTTCTCCAGCTAGAGGGTGGAAAGATACGAAAGAGAGCTCTCAAGTCTTGA 10185
Qy 3520 ATACTCATATGATGATGTC 3539
Db 10184 ACAGTGTGAGTACTGTCTC 10165

RESULT 8
AY19896 341 bp DNA linear PLN 24-JAN-2003
LOCUS Arabidopsis thaliana sequence flanking Ds3 end of Ds-EnhancerTrap
DEFINITION insertion from line ET7222.
ACCESSION AY19896
VERSION AY19896.1 GI:27895850

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS 1 (bases 1 to 341)
May, B.P., Simorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
TITLE Arabidopsis genomic sequences flanking Ds enhancer and gene traps
in transgenic lines
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 341)
May, B.P., Simorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R.,
McCombie, W.R. and Martienssen, R.A.
AUTHORS Direct Submission
TITLE Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown
JOURNAL Road, Cold Spring Harbor, NY 11724, USA
COMMENT http://genetrap.cshl.org.
FEATURES
Location/Qualifiers
1..341
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/clone="ET7222.Ds3.05.2002.jw20.341"
/ecotype="Landsberg"
/note="transgenic line ET7222"

ORIGIN

Query Match 8.3%; Score 304.4; DB 8; Length 341;
Best Local Similarity 99.4%; Pred. No. 3.3e-61;
Matches 316; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 497 ACAGTGGATTTGGTCTCTCCCATTTGCAATATGACGATTCACGCGGACGACCAAGC 556
Db 2 ACAGTGGATTTGGTCTCTCCCATTTGCAATATGACGATTCACGCGGACGACCAAGC 60
Qy 557 TCCGACGTAGCAACACCTCTACAACTATCTCTCCGACGAGAAATGGGCGACCGTC 616
Db 61 TCCGACGTAGCAACACCTCTACAACTATCTCTCCGACGAGAAATGGGCGACCGTC 120
Qy 617 TTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTGCGCACCGCCACCA 676

Db 121 TTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTGCGCACCGCCACCA 180
Qy 677 CCAACGGCACTCTGCTCTCTCTGCAACATCTATGATCTGTCCGAAACGGCAGCTCCCA 736
Db 181 CCAACGGCACTCTGCTCTCTCTGCAACATCTATGATCTGTCCGAAACGGCAGCTCCCA 240
Qy 737 TCCCATTTGATTTCTACCAAGTATTTAGAGCTCAACACATTTCTTAAACGATGAAATCA 796
Db 241 TCCCATTTGATTTCTACCAAGTATTTAGAGCTCAACACATTTCTTAAACGATGAAATCA 300
Qy 797 GAAGAGCATTCGAAGCTA 814
Db 301 GAAGAGCATTCGAAGCTA 318

RESULT 9
166494 7218 bp DNA linear PAT 28-DEC-1997
LOCUS Sequence 14 from Patent US 5670367.
ACCESSION 166494
VERSION 166494.1 GI:2724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner, F., Schefflinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
Location/Qualifiers
1..7218
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 1.6%; Score 57.6; DB 6; Length 7218;
Best Local Similarity 3.9%; Pred. No. 0.019;
Matches 15; Conservative 218; Mismatches 147; Indels 0; Gaps 0;
Qy 510 TCCTCTCCCATTCATATGACGATTCACGCGGACGACCAAACTCCGAGCTAGCCA 569
Db 1078 YY 1137
Qy 570 CAACACCTCTACACTATGCTGTCCGACGACCAATGGGCGACCGCTTCTCTCCGACTT 629
Db 1138 YY 1197
Qy 630 CAATTCACCTCCGATTCCT 689
Db 1198 YY 1257
Qy 690 CGTCTCTGCGACCATCTATGATCGTCCGAAACGACGATCCCATTCGATTTT 749
Db 1258 YY 1317
Qy 750 CTACCAAGTATTTAGAGCTCAACACATTTCTTACCGATGAAATCAAGAGCATTCGA 809
Db 1318 YY 1377
Qy 810 AGCTAGGTTTCGAACCGCGCATTCGCTTTCAGGAGAGAGCGCTTATACGCCGAG 869
Db 1378 YY 1437
Qy 870 ACAGATTCCTCAAGCTCTT 889
Db 1438 CCAATTCCTCTATCTCTT 1457
RESULT 10
AX083744 1141 bp DNA linear PAT 28-FEB-2001
LOCUS Sequence 22 from Patent WO011061.
ACCESSION AX083744

VERSION	AX083744.1	GI:13185472
KEYWORDS	synthetic construct	
SOURCE	synthetic construct	
ORGANISM	other sequences; artificial sequence.	
REFERENCE	1. Kunst, L. and Clemens, S.	
AUTHORS	Regulation of embryonic transcription in plants	
TITLE	Patent: WO 011061-A 22 15-FEB-2001;	
JOURNAL	UNIVERSITY OF BRITISH COLUMBIA (CA)	
FEATURES	Location/Qualifiers	
source	1..1141	
	/organism="synthetic construct"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:32630"	
	1..1141	
	/note="consensus sequence of A.t., L.a., and B.n. PAB1 promoters"	
ORIGIN		
Query Match	1.5%;	Score 54; DB 6; Length 1141;
Best Local Similarity	10.0%;	Pred. No. 0.12;
Matches	79; Conservative	311; Mismatches 389; Indels 8; Gaps 2;
QY	2393 TTAGCTGATGATCCCTGTTGGTAAACAATGAGCGCGGAGTGGAGCGCTGCTCTTATG 2455	
DB	67 KYRRATNNKSRWKKWKKYBVCANNYSBRHARMRDMKTAVBMTITNKGKGMHR 126	
QY	2453 CAGAAGCTGTAAGACCCCTGGA- AACTTGGAACTAATGATTAATGCAATTCGAGCTGG 2511	
DB	127 YWRRAEMDITDHNVTYANNAAWTTCCMDKDDKTRWKKNNATGMDDDTKHMMNN 186	
QY	2512 GTCTCAGAAGTAGCGGTGATGAAACTACTGTTGAAATGTCGGTTCGATATGTTAAAG 2571	
DB	187 NGCBVTVMWVYKTDRODSBKMYGMBMKWMSYDVYVWVWMDCKRKRVARVTRG 246	
QY	2572 GAGGCAAGTGTGAAGATCTGACCTGCTGCTGCTGGGCAATTGCACTGATTTCACTGTTCAGC 2633	
DB	247 RMRVYVMBABHARRRRYNNGWTBAAVYRRMTNNNNNNAKMKCRKKYMGWRABVNSTC 306	
QY	2632 CAGAAGTATTTCTTAAAGAGAGCTCATCTTTTCAACGCAAGATATGTTCTTATG 2691	
DB	307 TTWKSATTKVATSCVANNCRAGDANKHKMKMSAMGVYNNNNNNNNNTTAKAHBAW 366	
QY	2692 GAATCTGATGCTGCTACCATAGTATGTAATGATGCAATTTTCATATCTGCATTTG 2751	
DB	367 DMVWHSAAKKMHANAHAHYSRKKWTBYRKKTYVNNNNNGTTMKRMMAMVTKKMDMBGTYN 426	
QY	2752 CTCAAATATGCTGTTGTTGGAGTAAAGACATATGTTCCACITTAATACATGTCSSAA 2811	
DB	427 NNNNGGRTYYGMYTKKKMKMYUYKWRANNCKRBAHDHKTCTHNNITWYDKTYMNNCYKS 486	
QY	2812 AGTGTACCAAGATTAACAAGTGTGAGTAATTTCACTAATATGCT-----GCTT 2866	
DB	487 MTNGSHRBAALVYVYWMWRRYVYAHANNMMWMDYWKACMYKYBVCGRKMNNTAAWYTK 546	
QY	2865 GAATTTTGTGATCAACTGTAAGACAGAAATGTAAATTTCACTGCAATTTCTGTAG 2924	
DB	547 SSMWTSTSYRKKTKNSWRKSDTRSMGRANVYADAHYGYKMTTRMBWBSHTBHDAG 606	
QY	2925 AATAACGTAGATTAAGAGATGCTTAACTAGTGGCTTTGTCCAATTTTCTTCTGATT 2984	
DB	607 AAHYWMBWVBAKCHCMKAWYAKKKAAGAGGSSNNNNNNNNNNNNNNNATCARDDYAA 666	
QY	2985 TTTTCTTTTCGATTTAGGCTCAGTCAAGCTCAGACGATTCAGAGCACTTCCAGAAATG 3044	
DB	667 RMYAANAANKWYUYKAAANNAAYTTAHANNWGCANNATDTRRTWKKNNNNNAGTWKNNNN 726	
QY	3045 ATGCTAGAGCTGCAGAGATATATGATGCAGAGTGCAGAGATTAAGCTGCTGCTTTG 3104	
DB	727 NAKNAAAKYTAALAAVVKAAKKHMRANMKMARGMHDAABTTDGRNGATTKYTTNNN 786	
QY	3105 GGCTGATCACCGCATYGAATGTTTACAGAGGTGAGGAATTAATCTCAATTCATCATC 3164	

Dn		: : :	:	:	: : : :	:	:				
Oy	787 NTYRGVTTNRAARGMANNNNNNNNNNNNNGSGDMWVTWAAYANTGTNNNNNNNNNAAX							846			
Dd	947 AWWTNKW 853										
	3165 ATTGTCG T 3171 :::										
RESULT_11	AEO14820										
LOCUS	AEO14820	250029 bp	DNA	linear	INV 07-OCT-2002						
DEFINITION	Plasmodium falciparum 3D7 chromosome 14 section 5 of 13 of the complete sequence.										
VERSION	AEO14820 AEO14187										
KEYWORDS	AEO14820.1 GI:23497305										
ORGANISM	Plasmodium falciparum 3D7 Plasmodium falciparum 3D7 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.										
REFERENCE	1 (bases 1 to 250029) Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W., Carlton,J.M., Pail,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyes,S., Chan,W.-S., Nene,V., Shallow,S.J., Sub,B., Peterson,T., Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Garucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrett,B.										
TITLE	Genome sequence of the human malaria parasite Plasmodium falciparum										
JOURNAL	Nature 419 (6906), 498-511 (2002)										
PUBMED	12368864										
REFERENCE	2 (bases 1 to 250029) Gardner,M.J.										
AUTHORS	Gardner,M.J.										
TITLE	Direct Submission										
JOURNAL	Submitted (13-Sep-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA Location/Qualifiers										
FEATURES	source	1..250029	/organism=	"Plasmodium falciparum 3D7"							
			/mol_type=	"genomic DNA"							
			/isolate=	"J3D7"							
			/db_xref=	"taxon:36329"							
			/chromosome=	"14"							
				83..122							
	repeat_region		/rpt_type=tandem								
				307..346							
	repeat_region		/rpt_type=tandem								
				346..442							
	repeat_region		/rpt_type=tandem								
				575..646							
	repeat_region		/rpt_type=tandem								
				<833..>>3733							
	gene		/locus_tag=	"PF14_0238"							
	mRNA			<833..>>3733							
			/locus_tag=	"PF14_0238"							
	CDS			833..3733							
			/locus_tag=	"PF14_0238"							
			/codon_start=	1							
			/product=	"hypothetical protein"							
			/protein_id=	"AAM36851.1"							
			/db_xref=	"GI:23497306"							
			/translation=	"MNNDINNELINNLPFKFSIDISLSPKKCKDYIIPDEGGHTTWPBPSRPIGNVKKKEOKKKONEDSDDNDDNNNNSSDHNNNNSSDHNNNKRLVLILKETLKNTTYAGKLAKEMEEHYTDLLKKKSQAFPVLYLBKDGRKNSTPTIKWK EKICNIIEKAQQKKQNQLKKUNPKKHANI SPFDITTKYSLSLNDFNISDLSN LKADI INNEBERYNLTNSLFQSFMEOULPFTKRKEKKODYDNNVCREYEDEVKVPKG ITHIELLEGESEEGENKKKKKKKTLDIVHVNINKSLSYDNNVI PYEEDDKNTKI NNENNNSNPFEEDKELDIFSSPDJI PFNHKKKIKLIDNVNLKSLSYDVENVDSLII ASOQRKKCKKERBKPOKKKKKKKNNONSANEHNKKIKNNINDFNMMOSILENYEVY RKMPPECKTEBNVEKEIKRIUVRPLGYTSMBESDSHI KLCLENDDPDDDDNNDDND NDNDNDNDND DDNDNDNDGDNDNNNDNNNDNDNNKDI TNMLIINEBEERKONTLENLKNOIN							

```

IPTSKEIYYTKYSKKLDCNDLNKLSDEKETTKNISNDLELIKGEKKKNDPTFPI
SHNEKDKNYLFPYDHKGDDNYNDEKNNLPCHH CDNYNDEKNNLPCHH CDN
YNNEDLPFLSLMSGIYDKRKGIKNMECKSEICVDKDKLSINYDENNNKLLSYNIN
VODENDKLYKNEQIKODIYNDNENDKLYKNEQIKODIYNDNENDKLYKNDOKIKOD
IYNDNENDKLYKNDOKIKODIYNDNENDKLYKNEQIKODIYNDNENDKLYKNDOKIKOD
BEIDKREKSNICICIKYIDKIDMKULIDEIKKKKKKNVENSNSYSSEDDICFYEIYKF
ICSNINIEGNENIFDDTNINVENSNIDIMPOSIFNNQIKYLDNLISQIFIS"

repeat_region 1704..1740
/rpt_type=tandem
2309..2446
/rpt_type=tandem
2310..2443
/rpt_type=tandem
2313..2449
/rpt_type=tandem
2718..2893
/rpt_type=tandem
3065..3311
/rpt_type=tandem
3065..3273
/rpt_type=tandem
3776..3805
/rpt_type=tandem
4039..4094
/rpt_type=tandem
4260..4295
/rpt_type=tandem
4402..4439
/rpt_type=tandem
complement(<4519..>6088)
/locus_tag="PF14_0239"
complement(join(<4519..4588,4895..5573,5806..>6088))
/locus_tag="PF14_0239"
complement(join(4519..4588,4895..5573,5806..6088))
/locus_tag="PF14_0239"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN36852.1"
/db_xref="GI:23497307"
/translation="MEENREIEKTVSESEINNEYEELPLFASGNDLVKESLRI
ILSLAEKERIDHITIKNNKYIKTLISLSNSRCLLLECLINLSAQLPKELTERNI
EILFDMKREKEDNCIDYIMISNLTCKEYVYVLIDNDSNINIKEDNPKYF
PLNKLLYFPLPIKPSINKLSDRYIVSHVLNIISSIKSYVFKVAFLNKISDGI
LNVFRPAILPFIINLCINIEVIHPIFDHCYLPYVLSYLYNDVNTYSGSNNN
NNEINONTIHHITIMKSSILVSCVSKSRILILFYLCNRDYSRNDIENVTNLIEG
SNREAIATAT"

repeat_region 4677..4714
/rpt_type=tandem
5600..5641
/rpt_type=tandem
5745..5791
/rpt_type=tandem
6108..6152
/rpt_type=tandem
6168..6222
/rpt_type=tandem
6220..6268
/rpt_type=tandem
6420..6481
/rpt_type=tandem
6563..6621
/rpt_type=tandem
6718..6780
/rpt_type=tandem
6719..6790
/rpt_type=tandem
7543..7568
/rpt_type=tandem
7761..7816
/rpt_type=tandem
7824..7849
/rpt_type=tandem
7871..7930

```

```

gene /rpt_type=tandem
complement(<7953..>8438)
/locus_tag="PF14_0240"
mRNA complement(<7953..>8438)
/locus_tag="PF14_0240"
CDS complement(7953..8438)
/locus_tag="PF14_0240"
/codon_start=1
/product="ribosomal protein l21e, putative"
/protein_id="AAN36853.1"
/db_xref="GI:23497308"
/translation="MGKSGRSGTRRYKSKKRRHGECTANKYLEKNGYVDIV
CDSIQGKMPNRYHKGITGKIPIITRGGVVLVNRKVRHIEKQVCVREHVRKRC
NEDFLRKIKNAELIEAKLIKNNHINIKRTGSPKPAATIKVPPSKIITIEPLPIEE
Y"

repeat_region 8467..8550
/rpt_type=tandem
8473..8539
/rpt_type=tandem
8477..8537
/rpt_type=tandem
8723..8844
/rpt_type=tandem
8723..8840
/rpt_type=tandem
8782..8844
/rpt_type=tandem
9019..9059
/rpt_type=tandem
9019..9058
/rpt_type=tandem
9133..9189
/rpt_type=tandem
9221..9254
/rpt_type=tandem
9604..9637
/rpt_type=tandem
9668..9723
/rpt_type=tandem
10036..10073
/rpt_type=tandem
10445..10472
/rpt_type=tandem
10511..10551
/rpt_type=tandem
complement(<10565..>11080)
/locus_tag="PF14_0241"
complement(<10565..>11080)
/locus_tag="PF14_0241"
/locus_tag="PF14_0241"
/codon_start=1
/product="basic transcription factor 3b, putative"
/protein_id="AAN36854.1"
/db_xref="GI:23497309"
/translation="MEKISPEILAAAPAKLEKMGGNLRQIGKGSARRKIKYKHNSI
SNEKKNITLIKIGAFGVDIEICVYRTGDTLEFPKPKLCASLOSNTIIVGKFE
HKIDINKLFEGLKGNKYLDMNLLEKIKNDPINIKNINKSGDTPKREEDQANDVD
LVENPEVSKKE"

repeat_region 11429..11458
/rpt_type=tandem
11514..11584
/rpt_type=tandem
11514..11579
/rpt_type=tandem

Query Match 1.54; Score 54; DB 3; Length 250029;
Best Local Similarity 46.7%; Pred. No. 0.16;
Matches 171; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

Cy 11 TAAAGAGATACATTAAGCAATTGCTGATTTCAACAGATTTGCTGGCTATA 70
Db 218336 TAAAGAGATATGATGTGTAGATTATCTTGTGTGCAAGATTTATTTATTTGAA 218395

```

```

QY 71 GGATTCATGGCTCTGTTTGGCTTTACATTTACATGATTAAGTTTCGAAATTTACAC 130
DB 218396 TTGAAAAGTGTATTATATGATATGCTTTATTTGATATAGCTTAATTTTCTT 218455
QY 131 ATTTCAGTGGATGTTAAGAAAAGAGAGGAAATGATGGGTTTGGGGTTAACTTT 190
DB 218456 AATCTCGATGATCATTTGATGATTAATTAATTAATTAATTAATTAATTAATTA 218515
QY 191 AAGTATCAAGATTAAGTATGATGTTTACTGTTGCTCTATATGTAATTAAGAGCA 250
DB 218516 ATATTAATTAAGAAATATTTTGTGCTTAAGATTAATTAATTTCTTTAAATGTATCAT 218575
QY 251 ACTTCAACGGTTCTTAGTGATGATATTAATTAAGCATTTAATCAATCAATCAATCCGTG 310
DB 218576 ACATCGCATTTGCGAATTAAGACATATCTTTCTTTAATATTTATTTCTTTAAATCT 218635
QY 311 GCGACTGTAAATCATATGATGTTTATTTTATTTTTCAGTAGCTGTGATGTTTGTAT 370
DB 218636 ATTCCTGATTAACCATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 218695
QY 371 TTAATCT 376
DB 218696 TTAATCT 218701

RESULT 12
AL713891 205796 bp DNA linear HTG 30-MAY-2002
LOCUS Mus musculus chromosome 11 clone RP23-435M14, WORKING DRAFT
DEFINITION
SEQUENCE.
AL713891
AL713891.7 GI:21261921
HTG; HTGS PHASE2; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 205796)
Blakey, S.
Direct Submission
Submitted (27-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 29, 2002 this sequence version replaced gi:20067290.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: Dm435M14
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; Big Dye; 84% of reads
Consensus quality: 201571 bases at least Q40
Consensus quality: 203010 bases at least Q30
Consensus quality: 203553 bases at least Q20
Insert size: 205796; sum-of-contigs
Insert size: 201568; 3.2% error; agarose-fp
Quality coverage: 5.53x in Q20 bases; sum-of-contigs Quality
Coverage: 5.69x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 205796: contig of 205796 bp in length.

```

```

FEATURES
source
location/Qualifiers
1..205796
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-435M14"
/clone_1db="RP23-435M14"
/clone_1db="RP23-435M14"
1..205796
/feature="assembly_fragment:02806"

ORIGIN
Query Match 1.5%; Score 53.8; DB 2; Length 205796;
Best Local Similarity 56.5%; Pred. No. 0.18;
Matches 100; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 573 CACCTTACAACTATCTGCTCCGACGAAATGAGCGGACCGCTTCTTCCGACTTCA 632
DB 97775 CCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 97834
QY 633 TTTCACCTCGATTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 97894
DB 97835 TTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 97894
QY 693 CTCTCTGCAACATCTATGATGCTCCGAAAGCGACGATCCCATCCCATTTGATT 749
DB 97895 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 97951

RESULT 13
AL603913 215524 bp DNA linear ROD 16-FEB-2002
LOCUS Mouse DNA sequence from clone RP23-103H9 on chromosome 11, complete
DEFINITION
SEQUENCE.
AL603913
AL603913.14 GI:18855213
HTG.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 215524)
Heath, P.
Direct Submission
Submitted (16-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 21, 2002 this sequence version replaced gi:18476748.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WormPep; Information on the WormPep
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-103H9 is
from the RPI-23 Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
This sequence is the entire insert of clone RP23-103H9 The true
left end of clone RP23-298N21 is at 159538 in this sequence.
Location/Qualifiers

```

FEATURES

[illegible]

	repeat_region	26455..26494	/rpt type=tandem
	repeat_region	26541..26574	/rpt_type=tandem
	repeat_region	26638..26753	/rpt_type=tandem
	repeat_region	27339..27375	/rpt_type=tandem
	repeat_region	28248..28335	/rpt_type=tandem
	gene		/rpt_type=tandem
mRNA		<28491..>36165	/locus_tag="PF10_0001"
CDS		/note="Pf EMP1"	
		join(<28491..33893,34843..>36165)	/locus_tag="PF10_0001"
		join(28491..33893,34843..36165)	/locus_tag="PF10_0001"
		/codon_start=1	
		/product="erythrocyte membrane protein 1 (PEMP1)"	
		/protein_id="FANJ35201.1"	
		/db_xref="GI:23494866"	
		/translation="MAQSSGGGCGCEEDKAKTFMRDRIQKEVHDEKNAAVFQDY LNKLNTISITFGEETVAFTDPCELIEKRDELAAARDPCGKKEHFSEKRGACDD LKLEGNVRSGKGSKAGACAPYRLSLCKNFPNNMSKDLAKNDLLVDVCYAARY EASLILPYHDVLATYGDSQICVTIARSPIDIGIVRGLPIGHNRKKLEREGOL MEGKITHSNAKAISVLSKEIRREYRNALNDQWTKALTCDPENKLGSNALFHACRSER NGCSQAHKRCRPMTSDGKPNDQVPFYFDYFQYLWFEMMEDFCRKKIYDIVIK	

Query Match 1.5%; Score 53.8; DB 3; Length 250078;

Best Local Similarity 45.7%; Pred. No. 0.18;

Matches 187; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

Qy	93	TTTACATTTCACGTCTCAATAAATGTTTGAAATTTTACAATTCCAGTTGATGTTAAGAAA	152
Dd	36394	TGTACATAATATATCATTAATTAATATGTAATTAATATATATATATTAATATGATAACAAT	36455
Qy	153	AGAGAGGAATTGATGGCGGTTTTGTGGGTTTAAACTTAAAGTAGCAAGAATTAATGCA	212
Dd	36454	AATAGTTAAAAATTAATTAATTAATAAAATATTAACATTAATGGAACGATGAATATATATA	36511
Qy	213	TTGGTTTACGTGCTCTATATGTGTGTAATGAAATGGAAGCAATCCACCGTTCTTAGTGGA	272
Dd	36514	ATATATTTCATTTGCATTATTGTAATTAATTAATCAGCAATTAATATATATATATATAGA	3657
Qy	273	ATAGATTATTTTGAACGATTTAACATCATTAAGTCCGTGGCGACTGTAAATCATAGATTG	332
Dd	36574	AGTAGTAATTAACAGATACATTAATTAATTTTATATAGTAGTAACGATCTTTAAAAATTTA	3663
Qy	333	TTTTTATTTTTTTCAGTACGTGTGATGTTTTTTGATTTAACTTATATCAATCAAATTC	392
Dd	36634	ATATTTTATTTTTTATTTATATATTTTAACTTTTATTTATATTAATAAAAAAAAAAAAA	3669
Qy	393	AAATATCCATTAACCCCTAGACGACCAACAGCTCCTCAATATGTATTAACAGAACAAAGT	452
Dd	36694	AAATATATATATATATATTAATAATTTATTTAAATTTTTTTTCTTAATGTGTTTACATA	3675
Qy	453	TTTTGTAGTACCTTAATAAAGACATCCCATGAGAAAGCTGTAGTCAACGTC	501
Dd	36754	TAAATATATATAAAAAAAAAAAAAAAAAAAAAAAATTTAATAATATATC	36802

RESULT 15

AC116986 33321 bp DNA linear INV 12-MAR-2003

LOCUS Dictyostelium discoideum chromosome 2 map 2234041-2567370 strain

DEFINITION AX4, complete sequence.

ACCESSION AC116986 AC116425

VERSION AC116986.2 GI:28029247

KEYWORDS HTG.

SOURCE dictyostelium discoideum

ORGANISM Eukaryota; Mycetozoa; Dictyosteliiida; Dictyostelium.

REFERENCE 1 (bases 1 to 33321)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2005, 16:58:33 ; Search time 2053 Seconds
(without alignments)
11072.555 Million cell updates/sec

Title: US-10-600-070-3

Perfect score: 3667

Sequence: 1 tggctgcattaaagagagat.....ctataacataaggtctacaa 3667

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 309530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/prodata/1/pubpna/PC7_NEW_PUB.seq:*

3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/prodata/1/pubpna/PC7US_PUBCOMB.seq:*

7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*

14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*

18: /cgn2_6/prodata/1/pubpna/US10G_PUBCOMB.seq:*

19: /cgn2_6/prodata/1/pubpna/US10H_PUBCOMB.seq:*

20: /cgn2_6/prodata/1/pubpna/US10I_PUBCOMB.seq:*

21: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*

22: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

23: /cgn2_6/prodata/1/pubpna/US11A_PUBCOMB.seq:*

24: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*

25: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*

26: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3667	100.0	3667	19	US-10-600-070-3
2	3663.8	99.9	3667	19	US-10-600-070-10
3	1968.4	53.7	2679	20	US-10-739-930-227
4	1931	52.7	2637	19	US-10-600-070-130
5	1724	47.0	2406	19	US-10-600-070-1
6	1720.8	46.9	2406	19	US-10-600-070-9
7	1717.6	46.8	2406	19	US-10-600-070-128

Query Match	Score	DB	Length	App
Best Local Similarity	100.0%	Pred. No. 0;		
Matches	3667;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			

1 TGGTTCGATTAAAGAGATTAACATTTGCTTGAATTCACAGATTTTG 60

1 TGGTTCGATTAAAGAGATTAACATTTGCTTGAATTCACAGATTTTG 60

ALIGNMENTS

RESULT 1

US-10-600-070-3

Sequence 3, Application US/10600070

Publication No. US20040139500A1

GENERAL INFORMATION:

APPLICANT: Osteoryoung, Katherine W.

APPLICANT: Viltha, Stanislaw

APPLICANT: Koksharova, Olga A.

APPLICANT: Gao, Hong

TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of

TITLE OF INVENTION: Use

FILE REFERENCE: MSU-08153

CURRENT APPLICATION NUMBER: US/10/600, 070

CURRENT FILING DATE: 2003-06-20

NUMBER OF SEQ ID NOS: 206

SOFTWARE: PatentIn version 3.2

SEQ ID NO 3

LENGTH: 3667

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-10-600-070-3

QY 61 CTTGGCTAATAGGATTCATTCGCTCTGTTTGTCTTTTACATTTTACATGTCATATATAGTTTCG 120
DB 61 CTTGGCTAATAGGATTCATTCGCTCTGTTTGTCTTTTACATTTTACATGTCATATATAGTTTCG 120
QY 121 AATTTTACATTTTCAATTTGGAATGTTAAAGAAAGAGAGGGAATTTGAGGGGTTTGTGGG 180
DB 121 AATTTTACATTTTCAATTTGGAATGTTAAAGAAAGAGAGGGAATTTGAGGGGTTTGTGGG 180
QY 181 TTTTAACTTTAAAGTAGTCAAGAAATTAAGTCATTTGTTACTGTTGCTATATGTGTAA 240
DB 181 TTTTAACTTTAAAGTAGTCAAGAAATTAAGTCATTTGTTACTGTTGCTATATGTGTAA 240
QY 241 AATGAGGCAACTCCAAAGGTTCTTAAAGTAGAATTAATTTAGACGATTTAAATCAT 300
DB 241 AATGAGGCAACTCCAAAGGTTCTTAAAGTAGAATTAATTTAGACGATTTAAATCAT 300
QY 301 AAGTCCGTGGCCGCTGTAATCATATAGATTTGTTTATTTTATTTTGAAGCTGGTAT 360
DB 301 AAGTCCGTGGCCGCTGTAATCATATAGATTTGTTTATTTTATTTTGAAGCTGGTAT 360
QY 361 GTTTTGTATTTAACTTATCTCAAAATCAAAATTCATAAACCTTGAACGACCAAA 420
DB 361 GTTTTGTATTTAACTTATCTCAAAATCAAAATTCATAAACCTTGAACGACCAAA 420
QY 421 CAGTCTCTCAATATGTATAAAGAGAAAGATTTTGTAGTACCTTAAAAAGACATCCC 480
DB 421 CAGTCTCTCAATATGTATAAAGAGAAAGATTTTGTAGTACCTTAAAAAGACATCCC 480
QY 481 ATGGAAGCTGTAGTCAAGTCGCGCATTTGTCCTCCCAATTCGAATTAAGCCGATTAACA 540
DB 481 ATGGAAGCTGTAGTCAAGTCGCGCATTTGTCCTCCCAATTCGAATTAAGCCGATTAACA 540
QY 541 CCGGCGACGACAAAGCTCCGACGTAGCCACAAACCTCTCAACATATCTGCTCCGACG 600
DB 541 CCGGCGACGACAAAGCTCCGACGTAGCCACAAACCTCTCAACATATCTGCTCCGACG 600
QY 601 AAATGGGCGGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
DB 601 AAATGGGCGGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
QY 661 TTGGCCACCGGACGACCAACCGCATCTGCTCTCTGCGACCATCTTATTTGAATCCGCCC 720
DB 661 TTGGCCACCGGACGACCAACCGCATCTGCTCTCTGCGACCATCTTATTTGAATCCGCCC 720
QY 721 GAAAGCGACGTCGCCATCCCATTTGATTTTCAACAGGTAATGAGAGCTCAAAACATTTTC 780
DB 721 GAAAGCGACGTCGCCATCCCATTTGATTTTCAACAGGTAATGAGAGCTCAAAACATTTTC 780
QY 781 TTAACCGATGAATCAGAAAGCATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGGT 840
DB 781 TTAACCGATGAATCAGAAAGCATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGGT 840
QY 841 TTTCAGGACGACCGCTTAAATCAAGCGGACGACGATTTCTTCAAGTGTCTTCGAAACTCTG 900
DB 841 TTTCAGGACGACCGCTTAAATCAAGCGGACGACGATTTCTTCAAGTGTCTTCGAAACTCTG 900
QY 901 TCTAATCTCGGCTGTAAGAGAGTCAATGAAGTCTTCTTGAATGTAAGAGAGTCA 960
DB 901 TCTAATCTCGGCTGTAAGAGAGTCAATGAAGTCTTCTTGAATGTAAGAGAGTCA 960
QY 961 GTCATCACTGATGTTCTTGGGATTAAGGTAATTTGATTTTCGAAATTAAGATTTCTTC 1020
DB 961 GTCATCACTGATGTTCTTGGGATTAAGGTAATTTGATTTTCGAAATTAAGATTTCTTC 1020
QY 1021 GTTTTAAATTTCAATTTGATTAAGAGAGAACTTTATCTAGTGAAGTCTCTGGGGC 1080
DB 1021 GTTTTAAATTTCAATTTGATTAAGAGAGAACTTTATCTAGTGAAGTCTCTGGGGC 1080
QY 1081 TCTCTGTGTATTTGCAAGAGAGTGTGAGACTGATAGTCTTCTGAGGTTGGTGAAGCTCT 1140
DB 1081 TCTCTGTGTATTTGCAAGAGAGTGTGAGACTGATAGTCTTCTGAGGTTGGTGAAGCTCT 1140
QY 1141 GCTTAAGAGAGGTTGCTAAGTCTGTTTAAAGCAAGATGTGGTTTATGTTATGGCGCTTGC 1200

DB 1141 GCTTAAGAGAGGTTGCTAAGTCTGTTTAAAGCAAGATGTGGTTTATGTTATGGCGCTTGC 1200
QY 1201 GTTTTCAATGTCGAGGAGTGTATGCAATTTGATTCACCTGATTTTATTAATCTGCTTA 1260
DB 1201 GTTTTCAATGTCGAGGAGTGTATGCAATTTGATTCACCTGATTTTATTAATCTGCTTA 1260
QY 1261 TGAATTTGTGAAGAACTTTGAAGCTTTTACAGGTAATTTGAAGCTTTGATTTGATTTG 1320
DB 1261 TGAATTTGTGAAGAACTTTGAAGCTTTTACAGGTAATTTGAAGCTTTGATTTGATTTG 1320
QY 1321 ACGAGCTGTGGCTTTATTAAGAACTTTCTTGAATTTGATTTGATTTGATTTGATTTG 1380
DB 1321 ACGAGCTGTGGCTTTATTAAGAACTTTCTTGAATTTGATTTGATTTGATTTGATTTG 1380
QY 1381 GGAAGAGAGCAAGTATGCTGACACGGAATTAACGAGCAAAATGATGAGACTTTGGA 1440
DB 1381 GGAAGAGAGCAAGTATGCTGACACGGAATTAACGAGCAAAATGATGAGACTTTGGA 1440
QY 1441 AGAGATCACTCCGCTTATGCTTGAAGCTTATCCGCTTGTGATGATTAACGC 1500
DB 1441 AGAGATCACTCCGCTTATGCTTGAAGCTTATCCGCTTGTGATGATTAACGC 1500
QY 1501 TCGAAAAGACTTAAATGCTTTAAAGCGGTGCGGAATTTTGTGTCTGTGAGAGAGG 1560
DB 1501 TCGAAAAGACTTAAATGCTTTAAAGCGGTGCGGAATTTTGTGTCTGTGAGAGAGG 1560
QY 1561 TGAAGATCAAGCTCTTGTGGGGGTTTGAACCGTGAAGATTTATGATGAGAGGCTTTT 1620
DB 1561 TGAAGATCAAGCTCTTGTGGGGGTTTGAACCGTGAAGATTTATGATGAGAGGCTTTT 1620
QY 1621 ACGAATGACAGCTGCTGAGAGGTAATACGTTTGAATTTTATTTTCTTTAGCA 1680
DB 1621 ACGAATGACAGCTGCTGAGAGGTAATACGTTTGAATTTTATTTTCTTTAGCA 1680
QY 1681 TGAATTAATTTTATGCTTTTCAATTTTATGATTTGATTTGATTTGATTTGATTTG 1740
DB 1681 TGAATTAATTTTATGCTTTTCAATTTTATGATTTGATTTGATTTGATTTGATTTG 1740
QY 1741 GCTAACCCGAAGCAATTTTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACTTCTT 1800
DB 1741 GCTAACCCGAAGCAATTTTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACTTCTT 1800
QY 1801 GTGGCTCAAGCTTTTATGTTAAAGAGCAACCTTTTACAGATGCTGATTAAGCAATTC 1860
DB 1801 GTGGCTCAAGCTTTTATGTTAAAGAGCAACCTTTTACAGATGCTGATTAAGCAATTC 1860
QY 1861 CAGCAACTTCAAGAGCTTAAGGTAATGCTATGAGATTTCTGCAATGTTATGATTA 1920
DB 1861 CAGCAACTTCAAGAGCTTAAGGTAATGCTATGAGATTTCTGCAATGTTATGATTA 1920
QY 1921 CGGAATTAATTTGGAAGTATGCTGCTGTAAGAGGGAATCTGCTGCACTTATGAGC 1980
DB 1921 CGGAATTAATTTGGAAGTATGCTGCTGTAAGAGGGAATCTGCTGCACTTATGAGC 1980
QY 1981 AAAGTTGATGAATGCCGTAATGCTTGGGCTTGAACAGTGAAGATTTCAATATTAAGAT 2040
DB 1981 AAAGTTGATGAATGCCGTAATGCTTGGGCTTGAACAGTGAAGATTTCAATATTAAGAT 2040
QY 2041 CCAAGCTATTTGGAAGTGTGTTTGAAGATTTCAATCTGATGAATGATATCTCCCT 2100
DB 2041 CCAAGCTATTTGGAAGTGTGTTTGAAGATTTCAATCTGATGAATGATATCTCCCT 2100
QY 2101 GGAATTAATTTGGAAGTGTGTTTGAAGTGTGGAAGGAGGTTCTTCTCTAGGTTCAAGAC 2160
DB 2101 GGAATTAATTTGGAAGTGTGTTTGAAGTGTGGAAGGAGGTTCTTCTCTAGGTTCAAGAC 2160
QY 2161 ACCAAGATTAATTAATTTTAACTCGGGGACTATATGATGATGATTTGATTTGATTTG 2220
DB 2161 ACCAAGATTAATTAATTTTAACTCGGGGACTATATGATGATGATTTGATTTGATTTG 2220
QY 2221 TTGAAAAGAGTGAAGTATGTTCAAGGTTCTCTTTAGCTGCTGCACTATGAGCAAGG 2280

Db 2221 TTGGAAGAGTGAAGTAGTTCAGGGTTCCTTTAGCTGCTGCAACTATGCAAG 2280
 Qy 2281 ATGAGCCGAGCATGGAAGCTAGTCTATGAGGCACTGCAAGATTTTCTTCC 2340
 Db 2281 ATTGAGCCGAGCATGGAAGCTAGTCTATGAGGCACTGCAAGATTTTCTTCC 2340
 Qy 2341 CGCTATACAGATAGAACTCGGCTGAACCCAGGATGTCAGAGACAGTGTATAGTGA 2400
 Db 2341 CGCTATACAGATAGAACTCGGCTGAACCCAGGATGTCAGAGACAGTGTATAGTGA 2400
 Qy 2401 GATCTGTGTGTAACAAATGTAAGCCGTGATGTAAGCTGTGTATTTAGCAAGCT 2460
 Db 2401 GATCTGTGTGTAACAAATGTAAGCCGTGATGTAAGCTGTGTATTTAGCAAGCT 2460
 Qy 2461 GTAAGACCCCTGTAACCTTTGAACCTAATGATTAATGCAATTCAGCTGGGTCTCAGAG 2520
 Db 2461 GTAAGACCCCTGTAACCTTTGAACCTAATGATTAATGCAATTCAGCTGGGTCTCAGAG 2520
 Qy 2521 AGTAGCGTGAATGAATCTAGTGAATGTCGTCGTGATATGTTAAAGAGCAAGT 2580
 Db 2521 AGTAGCGTGAATGAATCTAGTGAATGTCGTCGTGATATGTTAAAGAGCAAGT 2580
 Qy 2581 GTGAAGATCCTAGCTGCTGTGTGTAAGCAATGTAAGCTGATTCAGTGTTCAGCCAGAGTAT 2640
 Db 2581 GTGAAGATCCTAGCTGCTGTGTGTAAGCAATGTAAGCTGATTCAGTGTTCAGCCAGAGTAT 2640
 Qy 2641 TTTCTTAAAGCACTCATCTTTCAACGCAAGATATGTTTCTTATAGATCTGAT 2700
 Db 2641 TTTCTTAAAGCACTCATCTTTCAACGCAAGATATGTTTCTTATAGATCTGAT 2700
 Qy 2701 GTGCTACATAGGATATGATTAATGATGCAATTTTCAATATCTGATCTCAATA 2760
 Db 2701 GTGCTACATAGGATATGATTAATGATGCAATTTTCAATATCTGATCTCAATA 2760
 Qy 2761 TGTCTTTTGTGAGCTAAGAACATAGTCTCCACTTAATACATGTCCTCAAAAGTTGACC 2820
 Db 2761 TGTCTTTTGTGAGCTAAGAACATAGTCTCCACTTAATACATGTCCTCAAAAGTTGACC 2820
 Qy 2821 AAGATTAAACAAGTGTGAGATTAATTTCACTAATATGCTGATTTTATGATCAA 2880
 Db 2821 AAGATTAAACAAGTGTGAGATTAATTTCACTAATATGCTGATTTTATGATCAA 2880
 Qy 2881 CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTTAAATACGTAGATAG 2940
 Db 2881 CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTTAAATACGTAGATAG 2940
 Qy 2941 AGATTGCTTAGTGTGCTTGTGTCACATTTCTTCTGTGATTTTCTTTCGANTT 3000
 Db 2941 AGATTGCTTAGTGTGCTTGTGTCACATTTCTTCTGTGATTTTCTTTCGANTT 3000
 Qy 3001 AGGCTCAGTCAAGCTGACATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGAGA 3060
 Db 3001 AGGCTCAGTCAAGCTGACATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTGAGA 3060
 Qy 3061 GAATATGATTCAGAGTGGGAGAGATTAAGTCTGCTTTGGGCTGATCCCGAT 3120
 Db 3061 GAATATGATTCAGAGTGGGAGAGATTAAGTCTGCTTTGGGCTGATCCCGAT 3120
 Qy 3121 AGAAATGTTACAGAGGTGAAGGAATTAATCTACATTCATCAATTTGTGAAAGCTGT 3180
 Db 3121 AGAAATGTTACAGAGGTGAAGGAATTAATCTACATTCATCAATTTGTGAAAGCTGT 3180
 Qy 3181 TGAACATGATTAATGCTGTGCTTGTGTTGATTTCTGTTATTAATGATTTGATGGGC 3240
 Db 3181 TGAACATGATTAATGCTGTGCTTGTGTTGATTTCTGTTATTAATGATTTGATGGGC 3240
 Qy 3241 GAATGCTGAAGATTGACCTGAGGCTTGTGATTTCTGTTATTAATGATTTGATGGGC 3300
 Db 3241 GAATGCTGAAGATTGACCTGAGGCTTGTGATTTCTGTTATTAATGATTTGATGGGC 3300
 Qy 3301 ATTATACACTGTTGAACTATCTGTTGACATGTTGACAGTCTCAGCAGATGGAACCCGTG 3360
 Db 3301 ATTATACACTGTTGAACTATCTGTTGACATGTTGACAGTCTCAGCAGATGGAACCCGTG 3360

Qy 3361 CTCTGTGAAGCACTCTGAGAGATCTGCTTCTATCTGATTTGTTGATTCAGAAA 3420
 Db 3361 CTCTGTGAAGCACTCTGAGAGATCTGCTTCTATCTGATTTGTTGATTCAGAAA 3420
 Qy 3421 ACAATGCTACTGATGTCAAGAACTTACACACACAGATACGAAATTTTCTGCTCAAGTACG 3480
 Db 3421 ACAATGCTACTGATGTCAAGAACTTACACACACAGATACGAAATTTTCTGCTCAAGTACG 3480
 Qy 3481 GGTGGAATACCTGAAGGCTCTGTTCTGATATGATTAATACATATGATGATGCT 3540
 Db 3481 GGTGGAATACCTGAAGGCTCTGTTCTGATATGATTAATACATATGATGATGCT 3540
 Qy 3541 GAGCTTCGAGATTTCTTGTGTTCTGTAATTTCTCTCTAAGTTAGTGTATTAATGA 3600
 Db 3541 GAGCTTCGAGATTTCTTGTGTTCTGTAATTTCTCTCTAAGTTAGTGTATTAATGA 3600
 Qy 3601 ACACAAAAAATTAAGCTTCTGACACACCTTTTCTGATCTTAACTATTAACATAAGG 3660
 Db 3601 ACACAAAAAATTAAGCTTCTGACACACCTTTTCTGATCTTAACTATTAACATAAGG 3660
 Qy 3661 GCTTACAA 3667
 Db 3661 GCTTACAA 3667

RESULT 2
 US-10-600-070-10
 ; Sequence 10, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Vitsha, Stanislaw
 ; APPLICANT: Kokscharova, Olga A.
 ; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; NUMBER OF SEQ ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 10
 ; LENGTH: 3667
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-600-070-10

Query Match 99.9%; Score 3663.8; DB 19; Length 3667;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3665; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGTTCGATTAAGAGAAATTAATTAAGCAATTTGCTTGAATTTCAACAAGATTTTG 60
 Db 1 TGTTCGATTAAGAGAAATTAATTAAGCAATTTGCTTGAATTTCAACAAGATTTTG 60
 Qy 61 CTGGCTATAGATTAATGCTGCTGTTGTTTAACTTAATTAATGATTAATGTTG 120
 Db 61 CTGGCTATAGATTAATGCTGCTGTTGTTTAACTTAATTAATGATTAATGTTG 120
 Qy 121 AATTTTACATTTCAAGTGTGATTTAAGAAAGAGGGAATTAATGATGAGGTTTGTGGG 180
 Db 121 AATTTTACATTTCAAGTGTGATTTAAGAAAGAGGGAATTAATGATGAGGTTTGTGGG 180
 Qy 181 TTTAACTTTAAGTAGTCAAGATTAAGTATGTTTACTGTTGCTATATAGTGA 240
 Db 181 TTTAACTTTAAGTAGTCAAGATTAAGTATGTTTACTGTTGCTATATAGTGA 240
 Qy 241 AATAAGCACTTCAACGCTTCTTAGTGAATTAATTAATTAAGAGATTAATCATCAT 300
 Db 241 AATAAGCACTTCAACGCTTCTTAGTGAATTAATTAATTAAGAGATTAATCATCAT 300
 Qy 301 AAAGTCGTGCGACTGTAACATCATAGATTTGTTTATTTTTCAGTAGCTGTGAT 360

Db	301	AAAGTCGGCGCATGTAACTATGATGATGTTTATTTTATTTTATGATAGCTGGAT	360
Qy	361	GTATTTTGAATTAATACTACTACTCAAAATCAAAATTCATAAACCTAGACGACAAA	420
Db	361	GTATTTTGAATTAATACTACTACTCAAAATCAAAATTCATAAACCTAGACGACAAA	420
Qy	421	CAGTCTCTCAATATGTAAACAGAACAAAGTTTGTATGAGCTTAAAGACACTCCC	480
Db	421	CAGTCTCTCAATATGTAAACAGAACAAAGTTTGTATGAGCTTAAAGACACTCCC	480
Qy	481	ATGGAAGCTCTGAGTCAAGTCGGCATTTGGTCTCTCCCAATTCGAATTAATGCCATTAACA	540
Db	481	ATGGAAGCTCTGAGTCAAGTCGGCATTTGGTCTCTCCCAATTCGAATTAATGCCATTAACA	540
Qy	541	CCGGCGACGACAAAGCTCCGACGAGGACAAACACTCTACACATCTGCTCCGACAGC	600
Db	541	CCGGCGACGACAAAGCTCCGACGAGGACAAACACTCTACACATCTGCTCCGACAGC	600
Qy	601	AAATGGGCGACGCTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC	660
Db	601	AAATGGGCGACGCTCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC	660
Qy	661	TTGGCCACCGCCACACACACGCGCACTCTGCTCTCTGCGACCATTTATGATCGTCCC	720
Db	661	TTGGCCACCGCCACACACACGCGCACTCTGCTCTCTGCGACCATTTATGATCGTCCC	720
Qy	721	GAAAGCCACGCTCCCATCCCATTTGATTTTCTACAGAGATTAAGAGCTCAAAACATTTT	780
Db	721	GAAAGCCACGCTCCCATCCCATTTGATTTTCTACAGAGATTAAGAGCTCAAAACATTTT	780
Qy	781	TTAAACGATGAATCAGAAAGACATTCGAACTTGGGTTTGGAAACGCGCAATTCGCT	840
Db	781	TTAAACGATGAATCAGAAAGACATTCGAACTTGGGTTTGGAAACGCGCAATTCGCT	840
Qy	841	TTTACGACACACGCTTTATATCAGCCGAGACAGTTCTTCAAGTCTCTGCGAACTCTG	900
Db	841	TTTACGACACACGCTTTATATCAGCCGAGACAGTTCTTCAAGTCTCTGCGAACTCTG	900
Qy	901	TCTAATCTCGGTCTAAGAGAGATTAAGATGAGTCTTCTGATGATGAAGAGCTACA	960
Db	901	TCTAATCTCGGTCTAAGAGAGATTAAGATGAGTCTTCTGATGATGAAGAGCTACA	960
Qy	961	GTCATCACTGATGTTCTTGGGATTAAGGTAATTCGATTTGGAATTAAGTTTCTTC	1020
Db	961	GTCATCACTGATGTTCTTGGGATTAAGGTAATTCGATTTGGAATTAAGTTTCTTC	1020
Qy	1021	GTATTTAATTTCAATGAATTTGAATTAAGAGAACTTTTATCTATGAGGTTCTGGGCG	1080
Db	1021	GTATTTAATTTCAATGAATTTGAATTAAGAGAACTTTTATCTATGAGGTTCTGGGCG	1080
Qy	1081	TCTCTGTATTTGCAAGAGAGTGGTGAAGCTGATGATGTTCTTGGGTTGGTGAAGCTCT	1140
Db	1081	TCTCTGTATTTGCAAGAGAGTGGTGAAGCTGATGATGTTCTTGGGTTGGTGAAGCTCT	1140
Qy	1141	GCTTAAGAGAGGTTGCTTAACTGTTTAAAGCAAGATGTGGTTTAACTTAACTGCTTC	1200
Db	1141	GCTTAAGAGAGGTTGCTTAACTGTTTAAAGCAAGATGTGGTTTAACTTAACTGCTTC	1200
Qy	1201	GTTTCTCGATGCTCTGAGGAGTGTATGCGATTCGATCTGATTTTATTAATCTGTTA	1260
Db	1201	GTTTCTCGATGCTCTGAGGAGTGTATGCGATTCGATCTGATTTTATTAATCTGTTA	1260
Qy	1261	TGAGTTTGTGAGAGACTTTGAAGCTTTTAAAGAGTGTGATCTTCTTGTGTAATTTG	1320
Db	1261	TGAGTTTGTGAGAGACTTTGAAGCTTTTAAAGAGTGTGATCTTCTTGTGTAATTTG	1320
Qy	1321	ACGAGCGTGGCTTTATTAAGAACTTTCTTGAATTTGATCTTGTATTAATGAGCTTGTGA	1380
Db	1321	ACGAGCGTGGCTTTATTAAGAACTTTCTTGAATTTGATCTTGTATTAATGAGCTTGTGA	1380
Qy	1381	GGAGGAAGAGCAATGATGCTTGAACCGGATTTACGTGCAAAATTAATGAGACTTTGGA	1440
Db	1381	GGAGGAAGAGCAATGATGCTTGAACCGGATTTACGTGCAAAATTAATGAGACTTTGGA	1440

Db	1381	GGAGGAAGAGCAATGATGCTTGAACCGGATTTACGTGCAAAATTAATGAGACTTTGGA	1440
Qy	1441	AGAGATCACTCCGCTTATGCTTGAAGCTTATGCTTGAAGCTTATGCTTGAAGCTTATG	1500
Db	1441	AGAGATCACTCCGCTTATGCTTGAAGCTTATGCTTGAAGCTTATGCTTGAAGCTTATG	1500
Qy	1501	TGCGAAAGACTTAAATGCTTTAAAGCGGTGTGCGGAATTTTGTGCTGTGTGAGAGAG	1560
Db	1501	TGCGAAAGACTTAAATGCTTTAAAGCGGTGTGCGGAATTTTGTGCTGTGTGAGAGAG	1560
Qy	1561	TGAGATCACTGCTTGTGCGGCTTGAACCGGATTTATGATGAGGCTTTT	1620
Db	1561	TGAGATCACTGCTTGTGCGGCTTGAACCGGATTTATGATGAGGCTTTT	1620
Qy	1621	ACGAATGACAGCTGCTGAGAGGATTAAGATTTAATCTTTTAAATTTCTTTAAGCA	1680
Db	1621	ACGAATGACAGCTGCTGAGAGGATTAAGATTTAATCTTTTAAATTTCTTTAAGCA	1680
Qy	1681	TGATTAATCTTTAAGCTTCTCATTTTAAATGATGTTGTGTGATGATTTTGTGA	1740
Db	1681	TGATTAATCTTTAAGCTTCTCATTTTAAATGATGTTGTGTGATGATTTTGTGA	1740
Qy	1741	GCTAACCCAGCAATTAATCCAGAGAGTCAATTTGAATTTAAGAGTTGACCTTCTT	1800
Db	1741	GCTAACCCAGCAATTAATCCAGAGAGTCAATTTGAATTTAAGAGTTGACCTTCTT	1800
Qy	1801	GTCGCTCAAGCTTTTATTTGTAAGAAAGCAACCTTTTAAAGAGATGCTGAATTAAG	1860
Db	1801	GTCGCTCAAGCTTTTATTTGTAAGAAAGCAACCTTTTAAAGAGATGCTGAATTAAG	1860
Qy	1861	CAGCACTTGAAGAGCTTAAAGTAAATGCTATGAGATTCCTGATGTTGATTAACA	1920
Db	1861	CAGCACTTGAAGAGCTTAAAGTAAATGCTATGAGATTCCTGATGTTGATTAACA	1920
Qy	1921	CGGAATTAATTTGGAAGATGATGCTTCCGCTTAAGAAAGGGAATCTGTGACCTGTAAT	1980
Db	1921	CGGAATTAATTTGGAAGATGATGCTTCCGCTTAAGAAAGGGAATCTGTGACCTGTAAT	1980
Qy	1981	AAAGTGAATGAATCCGATATGCTTGGGCTTAAGACAGTGAAGATTCACAAATTAAGAA	2040
Db	1981	AAAGTGAATGAATCCGATATGCTTGGGCTTAAGACAGTGAAGATTCACAAATTAAGAA	2040
Qy	2041	CCAGCTATTTGGAAGTGTGTTTGAAGATTCAAATCTGTATGAATGATATCTCCCT	2100
Db	2041	CCAGCTATTTGGAAGTGTGTTTGAAGATTCAAATCTGTATGAATGATATCTCCCT	2100
Qy	2101	GGAATTAAGCAATTTGTTGAAGACTGCTGTTGAGGAGGTTCTTCTCTAGCTTCAAGAC	2160
Db	2101	GGAATTAAGCAATTTGTTGAAGACTGCTGTTGAGGAGGTTCTTCTCTAGCTTCAAGAC	2160
Qy	2161	ACCAAGATTAAGAAATTTTAACTCCGAGGACTAATGATGATGATGATGATGATGATG	2220
Db	2161	ACCAAGATTAAGAAATTTTAACTCCGAGGACTAATGATGATGATGATGATGATGATG	2220
Qy	2221	TTGGAAGAGTGAAGTGTGAGGTTCTCTTTAAGCTGCTGCTGCTGCTGCTGCTGCTG	2280
Db	2221	TTGGAAGAGTGAAGTGTGAGGTTCTCTTTAAGCTGCTGCTGCTGCTGCTGCTGCTG	2280
Qy	2281	ATTGAGCCGAGATGGAAGCTGATGCTATGACAGGACTGAGAGAAAGTTTCTCTCC	2340
Db	2281	ATTGAGCCGAGATGGAAGCTGATGCTATGACAGGACTGAGAGAAAGTTTCTCTCC	2340
Qy	2341	CGCTATTAAGATTAAGAACTCCGCTGAACCCAGAGATGTGCAAGACAGTGTATGATGA	2400
Db	2341	CGCTATTAAGATTAAGAACTCCGCTGAACCCAGAGATGTGCAAGACAGTGTATGATGA	2400
Qy	2401	GATCTGTTGTTGAACATGTAAGCCGCTGATGCTGCTGCTGCTTATTTGCAAGAGCT	2460
Db	2401	GATCTGTTGTTGAACATGTAAGCCGCTGATGCTGCTGCTGCTTATTTGCAAGAGCT	2460
Qy	2461	GTAAGAGCCCTGGAAGAACTTTGAACATTAATGATTAAGCAATTTGAGCTGGGGTCTCAG	2520
Db	2461	GTAAGAGCCCTGGAAGAACTTTGAACATTAATGATTAAGCAATTTGAGCTGGGGTCTCAG	2520

QY 968 CTGATGTTCTTGGGATTAAGTAATTGCAATTCGAAATTAATAAGTTCTTGCTTTAA 1027
 |||||
 Db 602 CTGATGTTCTTGGGAT----- 618
 QY 1028 TTTTCATGATTTGATTAAGGAAGAACTTTATCTAGTGAAGTTCTCGGGGCTCTGCT 1087
 |||||
 Db 619 -----AAGTTCTCGGTCCTCTGT 639
 QY 1088 GTATTGCAAGAGGTGTGAGACTGAGATAGTTCTTCGGGTGGTGAAGCTCTGCTTAAG 1147
 |||||
 Db 640 GTATTGCAAGAGGTGTGAGACTGAGATAGTTCTTCGGGTGGTGAAGCTCTGCTTAAG 699
 QY 1148 GAAAGTTGCTTAAGTGTGTTAAGCAAGATGTGTTTAAAGGCTTGCGTTCTC 1207
 |||||
 Db 700 GAAAGTTGCTTAAGTGTGTTAAGCAAGATGTGTTTAAAGGCTTGCGTTCTC 759
 QY 1208 GATGTCGAGGGATGCTAATGGCAATGATGCACTGATTTTATTACTGGTTAGAGTTT 1267
 |||||
 Db 760 GATGTCGAGGGATGCTAATGGCAATGATGCACTGATTTTATTACTGGTTAGAGTTT 819
 QY 1268 GTTGAAGAACCTTTGAAGCTTTTAACAGTATGTTGACTTGCTTGTGTAAATTGAAGAAG 1327
 |||||
 Db 820 GTTGAAGAACCTTTGAAGCTTTTAC----- 844
 QY 1328 TTGGCTTTAATAAGAACTTTCTTGATTTTGATTACTTTGTTATTTAGCTTGTGTAGAGAA 1387
 |||||
 Db 845 -----AGAGGAA 852
 QY 1388 GGAGCAAGTAGCCTTGACCGGATTTTACGTGCACAAAATTGATGAGCTTTGGAAGATGC 1447
 |||||
 Db 853 GGAGCAAGTAGCCTTGACCGGATTTTACGTGCACAAAATTGATGAGCTTTGGAAGATGC 912
 QY 1448 ACTCCGGGTTATGTCCTTGAAGCTAATTGGCTTACCGCTTGATGATTAACGCTGCAAA 1507
 |||||
 Db 913 ACTCCGGGTTATGTCCTTGAAGCTAATTGGCTTACCGCTTGATGATTAACGCTGCAAA 972
 QY 1508 AGACTAATATGTTTAAGCGGTGTGCGGAATTTTGTGTGTGTGTGAGAGAGGTGAGCA 1567
 |||||
 Db 973 AGACTAATATGTTTAAGCGGTGTGCGGAATTTTGTGTGTGTGTGAGAGAGGTGAGCA 1032
 QY 1568 TCAGCTCTGTGTGGGGGTTTGAACCGTGAGAAAGTTTATGATGAGGGGCTTTTACGAATG 1627
 |||||
 Db 1033 TCAGCTCTGTGTGGGGGTTTGAACCGTGAGAAAGTTTATGATGAGGGGCTTTTACGAATG 1092
 QY 1628 ACAGCTGCTAGCAAGGTATACAGTTTAGAATCTTTTAAATTTCTTACATGATATTA 1687
 |||||
 Db 1093 ACAGCTGCTAGC----- 1105
 QY 1688 ACTTTAGGTTTCTCATTTTAATGATGTTGTGTGTGTGTGATCTTTTGTAGCTACCC 1747
 |||||
 Db 1106 -----AGTTGATCTTTTGTGTAGCTACCC 1129
 QY 1748 CAAGCAATATTTCCAGCAGAGCATTTGAAGTTTACGAAGTTGCACTTGCTCTTGAGCTC 1807
 |||||
 Db 1130 CAAGCAATATTTCCAGCAGAGCATTTGAAGTTTACGAAGTTGCACTTGCTCTTGAGCTC 1189
 QY 1808 AAGCTTTTATTTGTAAGAACCAACACCTTTTACAGATGCTGATTAAGCAATTCACAGAAC 1867
 |||||
 Db 1190 AAGCTTTTATTTGTAAGAACCAACACCTTTTACAGATGCTGATTAAGCAATTCACAGAAC 1249
 QY 1868 TTGACGAGGCTAAGGTATAGGCTATGAGAGATCTGCGATGTTGTATGATACACGGAAATA 1927
 |||||
 Db 1250 TTGACGAGGCTAAGGTATAGGCTATGAGAGATCTGCGATGTTGTATGATACACGGAAATA 1309
 QY 1928 ATTGGGAGATAGACTTCGGTCTAAGAAAGGGGAACTGTGCACTGCTTAATGGGCAAAAGTTG 1987
 |||||
 Db 1310 ATTGGGAGATAGACTTCGGTCTAAGAAAGGGGAACTGTGCACTGCTTAATGGGCAAAAGTTG 1369
 QY 1988 ATGAATCCGCTATGTGGTTGGGCTTACAGATGAGATTTCACAATATAGGAATCCAGCTTA 2047
 |||||
 Db 1370 ATGAATCCGCTATGTGGTTGGGCTTACAGATGAGATTTCACAATATAGGAATCCAGCTTA 1429
 QY 2048 TTGTGAGATTTGTTTGGAGAAATTCAAATGCTGATGACAAATGATGATCTCCCTGAGCTAT 2107

Db 1430 TTGTGAGATTTGTTTGGAGAAATTCAAATGCTGATGACAAATGATCTCCCTGAGCTAT 1489
 |||||
 QY 2108 GCAAAATTTGGAACCTGTGTGGCAGGGGTTGTCTTTCTTACGTTTCAGAGACACCAAG 2167
 |||||
 Db 1490 GCAAAATTTGGAACCTGTGTGGCAGGGGTTGTCTTTCTTACGTTTCAGAGACACCAAG 1549
 QY 2168 ATTAATAATTTTAACTCGGGGACCTATGATGATGCTTATGTTTGTAGATTACTTGGAAA 2227
 |||||
 Db 1550 ATTAATAATTTTAACTCGGGGACCTATGATGATGCTTATGTTTGTAGATTACTTGGAAA 1609
 QY 2228 GAGTGAAGTATGACGGGTTCTCCTTATGCTGTGCAACTATAGGCAAGATTTGAG 2287
 |||||
 Db 1610 GAGTGAAGTATGACGGGTTCTCCTTATGCTGTGCAAGCTATAGGCAAGATTTGAG 1669
 QY 2288 CCGAGCATGTGAAGCTAGTGTATGACAGGCACTGCAAGAAATTTTCTTCCGCTATA 2347
 |||||
 Db 1670 CCGAGCATGTGAAGCTAGTGTATGACAGGCACTGCAAGAAATTTTCTTCCGCTATA 1729
 QY 2348 CAGATGAACCTCGGCTGAACCAAGGATGTGACAGACAGTGTTTATGTTAGATCTTG 2407
 |||||
 Db 1730 CAGATGAACCTCGGCTGAACCAAGGATGTGACAGACAGTGTTTATGTTAGATCTTG 1789
 QY 2408 TTGGTAACTATGTAGCCGCTGATGTGAGCCTGCTGTATTTGCAAGCTGTAAGAC 2467
 |||||
 Db 1790 TTGGTAACTATGTAGCCGCTGATGTGAGCCTGCTGTCTTTATTTGCAAGCTGTAAGAC 1849
 QY 2468 CCTCTGAAAACCTTTGAAAACCTAATGATTAATGCAATTCGAGCTGGGGTCTCAGAGATGACG 2527
 |||||
 Db 1850 CCTCTGAAAACCTTTGAAAACCTAATGATTAATGCAATTCGAGCTGGGGTCTCAGAGATGACG 1909
 QY 2528 TTGATGAACCTACTGTTGAATATCGGTGCTGATATGTTTAAAGAGCAAGTGTGAAGA 2587
 |||||
 Db 1910 TTGATGAACCTACTGTTGAATATCGGTGCTGATATGTTTAAAGAGCAAGTGTGAAGA 1969
 QY 2588 TTCCTAGCTGTGGTGTGGCAATTTGACCTGATTCACGTTCACCGAAGTATTTTCTTA 2647
 |||||
 Db 1970 TTCCTAGCTGTGGTGTGGCAATTTGACCTGATTCACGTTCACCGAAGTATTTTCTTA 2029
 QY 2648 AAGACAGCTCATCTTTTCAACGCAAGATATGTTTTCTTATGAAATCTGATGTGCTTA 2707
 |||||
 Db 2030 AAGACAGCTCATCTTTTCAACGCAAGATATGTTTTCTTATGAAATCTGATGTGCTTA 2089
 QY 2708 CCAATAGATGATTAATATGATGCAATTTTCAATATCTGATGTGCTCAAAATATGCTTGT 2767
 |||||
 Db 2090 CCA----- 2092
 QY 2768 TTGTGAGCTAAGACATAGTTCCCACTTAATACATGTCCCAAAAGTTGTACCAAGATTA 2827
 |||||
 Db 2093 ----- 2092
 QY 2828 ACAAGTTGCTAGTAATAATTTCACTAATTAATGCTTGAAATTTTGTATCAAACTGTAGA 2887
 |||||
 Db 2093 ----- 2092
 QY 2888 CAGAAATGAATTTTCACTCTCAACATTTCTGTTTGAATTAACGTAGATTAAGATTGC 2947
 |||||
 Db 2093 ----- 2092
 QY 2948 CTTAGTGTGCTTTGTGCCAACTTTTCTTCTTGATTTTTCCTTTTCCATTTAGGGTCA 3007
 |||||
 Db 2093 -----TAGGGTCA 2100
 QY 3008 GTGAGAGCTGACGATTCAGAAAGCACTTCCAGATGATGATGCTAGGACTGAGAGAAATATA 3067
 |||||
 Db 2101 GTGAGAGCTGACGATTCAGAAAGCACTTCCAGATGATGATGCTAGGACTGAGAGAAATATA 2160
 QY 3068 GTATCCAAAGTGAGAGAAATTAATGCTGTGGCTTTTGGGCTGATCAACGCAATAGAAATG 3127
 |||||
 Db 2161 GTATCCAAAGTGAGAGAAATTAATGCTGTGGCTTTTGGGCTGATCAACGCAATAGAAATG 2220
 QY 3128 TTACGAGAGGTGAGGAATTAATCTACATTAATCAATTTGTGTGAAAACTGTTGAGACAT 3187
 |||||

Db	2221	TTACAG-----	2227
Qy	3188	GATTATAGCTGGTGCCCTGTTGATTCTGTATTATTAAGGTTTGGATGGCGAATGCT	3247
Db	2228	-----AGGTTTGGATGGCGAATGCT	2249
Qy	3248	GAAGATTGGACTGACAGAGCAGCTGAACTGGCAGCTTGGGTGGTTATGATTATAC	3307
Db	2250	GAAGATTGGACTGACAGAGCAGCTGAACTGGCAGCTTGGGTGGTTATGATTATAC	2309
Qy	3308	ACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCCGTGCTTGCT	3367
Db	2310	ACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCCGTGCTTGCT	2369
Qy	3368	GGAACCACTCTGGAGAGTGCTGCTGTATCTGATTTGATTCATCAGAAAACAATGC	3427
Db	2370	GGAGCAACTCTGGAGAGTCTGCTGTCTATCTGATTTGATTCATCAGAAAACAATGC	2429
Qy	3428	TACTGATGTGAGAACTTACACAAAGAATTCGAGTTTCTGTCACAGTCAAGGTGGAA	3487
Db	2430	TACTGATGTGAGAACTTACACAAAGAATTCGAGTTTCTGTCACAGTCAAGGTGGAA	2489
Qy	3488	AATGACTGAAGGCTCTGTTCTTGCAATCAATAATATACATATGATGAGATGCTGAAGCTTG	3547
Db	2490	AATGACTGAAGGCTCTGTTCTTGCAATCAATAATATACATATGATGAGATGCTGAAGCTTG	2549
Qy	3548	CGAGATTCCTTTGGTCTGTAAATTCCTCTCTAAGTTAGTGTATTAATGAACACAAA	3607
Db	2550	CGAGATTCCTTTGGTGTGTAATTCCTCTCTAAGTTAGTGTATTAATGAACACAAA	2609
Qy	3608	AAATATTAAGTCTTGACACACCCCTTTCTTGATCTAAACTATATACATAAGGCTACAA	3667
Db	2610	AAATATTAAGTCTTGACACAAAAAATCTTATCTTAAACTATATACATAAGGCTACAA	2669

Qy	548	CGAAGAGCTCCGACGTAAGCAACAACCTCTAACAATACTGCTCCGACGACAAATGGG	607
Dd	181	CGAAGAGCTCCGACGTAAGCAACAACCTCTAACAATACTGCTCCGACGACAAATGGG	240
Qy	608	CGAAGAGCTCTCTCCGACTTCAATTTCACTTCGATTCCTCCTCTCTCTTGGCCA	667
Dd	241	CGAAGAGCTCTCTCCGACTTCAATTTCACTTCGATTCCTCCTCTCTCTCTTGGCCA	300
Qy	668	CGGACACACACCGGCACTCTGCTCTGCGACAATACTATGATCGTCCGAAAGCC	727
Dd	301	CGGACACACACCGGCACTCTGCTCTGCGACAATACTATGATCGTCCGAAAGCC	360
Qy	728	ACGTCGCCATCCCAATTGATTTCTTACAGGTAATTAGAGCTCAACAACATTTCTTAAACG	787
Dd	361	ACGTCGCCATCCCAATTGATTTCTTACAGGTAATTAGAGCTCAACAACATTTCTTAAACG	420
Qy	788	ATGGAATTCAGAAAGCACTTCCGAACTGGGTTTGGAAACCGCGGCATTCGGTTTACGCG	847
Dd	421	ATGGAATTCAGAAAGCACTTCCGAACTGGGTTTGGAAACCGCGGCATTCGGTTTACGCG	480
Qy	848	ACGACGCTTTAATACGCGGAGACAGATTCTTCAAGCTGCTTGGAAACTCTGTCTATTC	907
Dd	481	ACGACGCTTTAATACGCGGAGACAGATTCTTCAAGCTGCTTGGAAACTCTGTCTATTC	540
Qy	908	CTCGGTCAGAGAGAGTACATGAAGCTCTTCTTGAATGATGAAGAGCTACAGTATCA	967
Dd	541	CTCGGTCAGAGAGAGTACATGAAGCTCTTCTTGAATGATGAAGAGCTACAGTATCA	600
Qy	968	CTGATGTTCTTGGGATTAAGGTATTTGATTTCCGAATAATAAATTTCTTCGTTTAA	1027
Dd	601	CTGATGTTCTTGGGATTT-----	617
Qy	1028	TTTCATAATTCGATTAAGAGAGCACTTTTATCTAGTAAGGTTCTCGGAGCTCTCTGT	1087

RESULT 4
 US-10-600-070-130
 , Sequence 130, Application US/10600070
 , Publication No. US20040139500A1
 , GENERAL INFORMATION:
 , APPLICANT: Ostryoung, Katherine W.
 , APPLICANT: Vitna, Stanislaw
 , APPLICANT: Kokeharova, Olga A.
 , APPLICANT: Gao, Hongbo
 , TITLE OR INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
 , FILE REFERENCE: MSU-08153
 , CURRENT APPLICATION NUMBER: US/10/600,070
 , CURRENT FILING DATE: 2003-06-20
 , NUMBER OF SEQ ID NOS: 206
 , SOFTWARE: PatentIn version 3.2
 , SEQ ID NO 130
 , LENGTH: 2637
 , TYPE: DNA
 , ORGANISM: Arabidopsis thaliana
 , US-10-600-070-130

Qy	1088	GTATTGCAAGAGGTGGTGAGACTGAGATGATTCTTGCGGTGGTGAGGCTGCTTAAG	1147
Db	639	GTATTGCAAGAGGTGGTGAGACTGAGATGATTCTTGCGGTGGTGAGGCTGCTTAAG	698
Qy	1148	GAGAGGTGCTTAAGTCGTTTAAGCAAGATGCGTTTAACTATATGCGCTTGCCTTTCTC	1207
Db	699	GAGAGGTGCTTAAGTCGTTTAAGCAAGATGCGTTTAACTATATGCGCTTGCCTTTCTC	758
Qy	1208	GATGCTCGAGGGAGTGCATAGGCACTGATGATCCACTGATTTTATTACTGGTATAGAGTTT	1267
Db	759	GATGCTCGAGGGAGTGCATAGGCACTGATGATCCACTGATTTTATTACTGGTATAGAGTTT	818
Qy	1268	GTTGAGGAGCTTTGAAGCTTTTACAGGTAGATTGACTTGCTTTGGTAAATTGACGACGC	1327
Db	819	GTTGAGGAGCTTTGAAGCTTTTAC-----	843
Qy	1328	TTGGCTTTATTAAGAACTTTCTTGATTGATACCTTTGTTATTTGAGCTCTTGCTAGAGGAA	1387
Db	844	-----ACGAGGAA	851

	Query Match	Similarity	Score	1931;	DB	19;	Length	2637;
	Best Local	Similarity	80.4%;	Pred.	No. 0;			
	Matches	2616;	Conservative	0;	Mismatches	5;	Indels	632;
							Gaps	5;
Qy	368	GATTTAACTTATCTACTCAAAATCAAAATTCATATAACCTTAGACGCAAAACAGTCTC						427
Db	1	GATTTAACTTATCTACTCAAAATCAAAATTCATATAACCTTAGACGCAAAACAGTCTC						60
Qy	428	TTCAATATGTAAACAGAACAAAGTTTGTAGTAGGCTTAAAAAGACATTCCTCATGAG						487
Db	61	TTCAATATGTAAACAGAACAAAGTTTGTAGTAGGCTTAAAAAGACATTCCTCATGAG						120
Qy	488	CTCTAGTCACGTCGGGCAATTGGTCTCTCCCATTTCCAAATATGCGCATTAACACCGGCGA						547
Db	121	CTCTAGTCACGTCGGGCAATTGGTCTCTCCCATTTCCAAATATGCGCATTAACACCGGCGA						180

Db	852	GGAGCAAGTACCTTGGACCCGGATTACGTGCACAAATTGATGAGACTTTGGAGAAGATC	911
Oy	1448	ACTCCGGCTTATGCTTGGAGCTACTTGGCTTACCGCTTGGTGA TGA TTACGCTGCCAAA	1507
Db	912	ACTCCGGCTTATGCTTGGAGCTACTTGGCTTACCGCTTGGTGA TGA TTACGCTGCCAAA	971
Oy	1508	AGACTAATGATTTAACCGGTGTGCGCAATATTTTGTGCTCTGTTGGAGAGGTGAGCA	1567
Db	972	AGACTAATGATTTAACCGGTGTGCGCAATATTTTGTGCTCTGTTGGAGAGGTGAGCA	1031
Oy	1568	TCAGCTCTTGTGGGGGTTTGA CCGCTGAGAGTTTATGAA TGA AGCGGTTTTTACGAATG	1627
Db	1032	TCAGCTCTTGTGGGGGTTTGA CCGCTGAGAGTTTATGAA TGA AGCGGTTTTTACGAATG	1091
Oy	1628	ACAGCTGCTGAGCAGTATACAGTTTAAATCCTTTTTTAAATTTCTTTAGCATGATATA	1687

TITLE OF INVENTION: Use
 FILE REFERENCE: MSU-08153
 CURRENT APPLICATION NUMBER: US/10/600.070
 CURRENT FILING DATE: 2003-06-20
 NUMBER OF SEQ ID NOS: 206
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 1
 LENGTH: 2406
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-10-600-070-1

Query Match 47.0% Score 1724; DB 19; Length 2406;
 Best Local Similarity 79.2%; Pred. No. 0;
 Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;

QY 481 ATGGAAGCTCTGATGATCGATGCGGATGATGCTCTCCCATTCACATTTATGCGGATTAACA 540
 DB 1 ATGGAAGCTCTGATGATCGATGCGGATGATGCTCTCCCATTCACATTTATGCGGATTAACA 60
 QY 541 CCGGCGACGACAAAGCTCCGACGTAAGCCACACACTCTACACTATCTGCTCGCCAGC 600
 DB 61 CCGGCGACGACAAAGCTCCGACGTAAGCCACACACTCTACACTATCTGCTCGCCAGC 120
 QY 601 AATAGGCGCGACCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTGCTCTCTCC 660
 DB 121 AATAGGCGCGACCGCTCTCTCTCCGACTTCAATTTCACTCCGATTCCTGCTCTCTCC 180
 QY 661 TTGCGCACCGCCACACACACGCGCATCTGCTCTCTGCGCACCATCTATTTGATCGTCCC 720
 DB 181 TTGCGCACCGCCACACACACGCGCATCTGCTCTCTGCGCACCATCTATTTGATCGTCCC 240
 QY 721 GAAGCGCAGTCCCGCATCCCATTTGATTCTACAGGATTTAGAGCTCAACACATTTTC 780
 DB 241 GAAGCGCAGTCCCGCATCCCATTTGATTCTACAGGATTTAGAGCTCAACACATTTTC 300
 QY 781 TTAAACGATGGAATTCAGAGAGCATTTGAGGATTCGAAACGCGCGCAATTCGGT 840
 DB 301 TTAAACGATGGAATTCAGAGAGCATTTGAGGATTCGAAACGCGCGCAATTCGGT 360
 QY 841 TTGACGACGACGCTTTAATCAAGCGGAGACAGATTTCTCAAGCTGCTTGCGAAATCTG 900
 DB 361 TTGACGACGACGCTTTAATCAAGCGGAGACAGATTTCTCAAGCTGCTTGCGAAATCTG 420
 QY 901 TCTAATCTCGGCTTGAAGAGATACATGAAGGCTCTTGTATGATGAAGAGCTACA 960
 DB 421 TCTAATCTCGGCTTGAAGAGATACATGAAGGCTCTTGTATGATGAAGAGCTACA 480
 QY 961 GTCATCTGATGCTTCTTGGAATGAAGTAAATTCGATTCGGAATAATAAGTTTCTTC 1020
 DB 481 GTCATCTGATGCTTCTTGGAATGAAGTAAATTCGATTCGGAATAATAAGTTTCTTC 504
 QY 1021 GTTTTAATTCATGAATTGATTAAGAAGAACTTTATCTAGTGAAGTTCTTGAGGC 1080
 DB 505 -----AAGTTCTTGAGGC 518
 QY 1081 TCTCTGTATTCAGAAAGGTGTGAGACTGAGATAGTTCTTGAGGTGTGAGGCTCT 1140
 DB 519 TCTCTGTATTCAGAAAGGTGTGAGACTGAGATAGTTCTTGAGGTGTGAGGCTCT 578
 QY 1141 GCTTAAGGAAGGTGTGAGGTGTGAGGTGTGAGGTGTGAGGTGTGAGGTGTGAGGTGTG 1200
 DB 579 GCTTAAGGAAGGTGTGAGGTGTGAGGTGTGAGGTGTGAGGTGTGAGGTGTGAGGTGTG 638
 QY 1201 GTTCTCGATGTCTGAGGAGTCTATGACATTTGATGATCCTGATTTTATCTGCTTA 1260
 DB 639 GTTCTCGATGTCTGAGGAGTCTATGACATTTGATGATCCTGATTTTATCTGCTTA 698
 QY 1261 TGAAGTTGTGAGGAGCTTTGAAGCTTTTACAGGATGTTTGAAGCTTTGCTTTGTAATTTG 1320
 DB 699 TGAAGTTGTGAGGAGCTTTGAAGCTTTTAC-----730
 QY 1321 ACGAGCGTTGGCTTTAATAAGAACTTCTGATTTGATTAATTTGATGATCTTTGCTTA 1380

DB 731 -----A 731
 QY 1381 GAGGAGAGGAGCAATGACCTTTCACCCGATTTACGTCACAAATTTGATGAGACTTTGCA 1440
 DB 732 GAGGAGAGGAGCAATGACCTTTCACCCGATTTACGTCACAAATTTGATGAGACTTTGCA 791
 QY 1441 AGAGATCACTCCGAGTATGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 1500
 DB 792 AGAGATCACTCCGAGTATGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 851
 QY 1501 TCGAAAAAGACTAATGCTTGAAGCGGTGTGCGAATATTTGTGCTGTGTGAGAGAG 1560
 DB 852 TCGAAAAAGACTAATGCTTGAAGCGGTGTGCGAATATTTGTGCTGTGTGAGAGAG 911
 QY 1561 TGAAGCATCAGCTCTTGTGTGGGGGTTCAGCCGTGAGAACTTTATGAATGAGGCTTTT 1620
 DB 912 TGAAGCATCAGCTCTTGTGTGGGGGTTCAGCCGTGAGAACTTTATGAATGAGGCTTTT 971
 QY 1621 ACGAATGACAGCTGTGAGCAGGATACAGTTTAGATACCTTTTAAATTTCTTTAGCA 1680
 DB 972 ACGAATGACAGCTGTGAGC-----991
 QY 1681 TGAATATACTTAGGTTTCTCATTTTAATGATATGTGTGTGATGATCTTTTGTGA 1740
 DB 992 -----AGTTGATCTTTTGTGA 1008
 QY 1741 GCTAACCCCAAGCAATATTCAGACAGACTATTGGAATTTACGAAGTTGCACTTGCTCTT 1800
 DB 1009 GCTAACCCCAAGCAATATTCAGACAGACTATTGGAATTTACGAAGTTGCACTTGCTCTT 1068
 QY 1801 GTGGCTCAAGCTTTTATGGAAGAGCAACCTTTTACAGATGCTGATTAAGCAATTC 1860
 DB 1069 GTGGCTCAAGCTTTTATGGAAGAGCAACCTTTTACAGATGCTGATTAAGCAATTC 1128
 QY 1861 CAGCAACTTCAGCAGCTAAGATTAATGCTATGAGATTCCTGCAATGTTGTATGATACA 1920
 DB 1129 CAGCAACTTCAGCAGCTAATGATTAATGCTATGAGATTCCTGCAATGTTGTATGATACA 1188
 QY 1921 CGGAATATTTGGAAGATTAATGCTGCTTGAAGAGGGAATCTTGTGCACTGCTTATAGGC 1980
 DB 1189 CGGAATATTTGGAAGATTAATGCTGCTTGAAGAGGGAATCTTGTGCACTGCTTATAGGC 1248
 QY 1981 AAGTTGATGAATGCGGCTAATGCTTGGGCTTGAAGAGTGAAGATTCACATATATAGGAT 2040
 DB 1249 AAGTTGATGAATGCGGCTAATGCTTGGGCTTGAAGAGTGAAGATTCACATATATAGGAT 1308
 QY 2041 CCAAGTATGAGGATTTGTTTGAAGATTTCAATGCTGATGACATATGATCTCCCT 2100
 DB 1309 CCAAGTATGAGGATTTGTTTGAAGATTTCAATGCTGATGACATATGATCTCCCT 1368
 QY 2101 GAGCTATGCAAAATTTGTGAAAACCTGCTGAGAGGGGTGCTTTCTTCTAGGTTCAAGAC 2160
 DB 1369 GAGCTATGCAAAATTTGTGAAAACCTGCTGAGAGGGGTGCTTTCTTCTAGGTTCAAGAC 1428
 QY 2161 ACCAATAATTAATACTTAACTCGGGGACATATGATGATGCTATGATTTGAGTTAC 2220
 DB 1429 ACCAATAATTAATACTTAACTCGGGGACATATGATGATGCTATGATTTGAGTTAC 1488
 QY 2221 TTGGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2280
 DB 1489 TTGGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1548
 QY 2281 ATTGAGCCGACGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2340
 DB 1549 ATTGAGCCGACGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1608
 QY 2341 CGCTATACATGGAATGCTGGGCTGAACCCAAGGATGGAAGGAGAGAGAGAGAGAGTGA 2400
 DB 1609 CGCTATACATGGAATGCTGGGCTGAACCCAAGGATGGAAGGAGAGAGAGAGAGAGTGA 1668
 QY 2401 GATCTGTGTGATACAAATGATAGCCGTGATGATGAGCTGGTGTCTTATTTGCAAGAGCT 2460

Db 1669 GATCCTGTTGGTAAACAATGTAAGCCGCTGATGCTGCTTTATTTGCAAGACT 1728
Qy 2461 GTAAGACCTCTGAAAACTTTGAAACTATGATTAATGCAATTCAGCTGGGGTCTCAGAG 2520
Db 1729 GTAAGACCTCTGAAAACTTTGAAACTATGATTAATGCAATTCAGCTGGGGTCTCAGAG 1788
Qy 2521 AGTAGCGTTGATGAAACTTACTGTTGAAATGTCCTGCTGATTAATGTTAAAGAGGCAAGT 2580
Db 1789 AGTAGCGTTGATGAAACTTACTGTTGAAATGTCCTGCTGATTAATGTTAAAGAGGCAAGT 1848
Qy 2581 GTGAAGATCCAGTGTGCTGGTGGCAATGGAATGCACTGATTCAGCTGTCAGCAGAGATAT 2640
Db 1849 GTGAAGATCCAGTGTGCTGGTGGCAATGGAATGCACTGATTCAGCTGTCAGCAGAGATAT 1908
Qy 2641 TTTCTTAAAAAGCAGCTCATCTTTTCAACGCAAGGATAATGTTTCTTATGAAATCGAT 2700
Db 1909 TTTCTTAAAAAGCAGCTCATCTTTTCAACGCAAGGATAATGTTTCTTATGAAATCGAT 1968
Qy 2701 GTGCGTACCAATAGATGATTAATGATGCAATTTTCAATATCTGCATTCGCAAAATA 2760
Db 1969 GTGCGTACCA----- 1978
Qy 2761 TGCTGTTTTGTGAGTAAAGACATAGTCCACTAATACATGTCGCAAAAGTTGATCC 2820
Db 1979 ----- 1978
Qy 2821 AAGATTAAACAAGTTGCTGATGAAATTTCACTAATTAATGCTGTAATTTTGTGATCAAA 2880
Db 1979 ----- 1978
Qy 2881 CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGAAATAGAGATTAAG 2940
Db 1979 ----- 1978
Qy 2941 AGATGCTTAAAGTGTGGCTTTGTCCAACTTTTCTTCTGATTTTCTTTTGCATTT 3000
Db 1979 ----- 1979
Qy 3001 AGGATCAGTCAAGCTGACGATTCAGAAAGCATTTCCAGAAATGATGCTAGAGCTGAGCA 3060
Db 1980 AGGATCAGTCAAGCTGACGATTCAGAAAGCATTTCCAGAAATGATGCTAGAGCTGAGCA 2039
Qy 3061 GAATATATGATTCAGAGTGGCAAGAAATTAAGTCTTGGCTTTGGGCTGATCACCGCAT 3120
Db 2040 GAATATATGATTCAGAGTGGCAAGAAATTAAGTCTTGGGCTGATCACCGCAT 2099
Qy 3121 AGAAATGTTTCAAGAGTGAAGGAAATTAATCTAATCAATCAATTTGTGAAAACTGT 3180
Db 2100 AGAAATGTTTCAAG----- 2113
Qy 3181 TGACATGATTAATAGTCTGGTCTTGTGATTCGTGTAATTAAGTTTTGGATGGGC 3240
Db 2114 ----- 2128
Qy 3241 GAATGCTGAAGATTTGGACTGACAGAGCAGCTGAACTGGCGAGCTTGGGTTTATG 3300
Db 2129 GAATGCTGAAGATTTGGACTGACAGAGCAGCTGAACTGGCGAGCTTGGGTTTATG 2188
Qy 3301 AATATACACGTGTTGAAACTATCTGTTGACAGTGAAGTCTCAGCAGATGAAACCCGTG 3360
Db 2189 AATATACACGTGTTGAAACTATCTGTTGACAGTGAAGTCTCAGCAGATGAAACCCGTG 2248
Qy 3361 CTCTGTTGAAGCAACTCTGAGAGTCTGCTGTCTATCTGATTTGGTTCAATCAGAAA 3420
Db 2249 CTCTGTTGAAGCAACTCTGAGAGTCTGCTGTCTATCTGATTTGGTTCAATCAGAAA 2308
Qy 3421 ACAATGCTACTGATGTCAGAACCTTACACAACAGATTCGAAGTTTCTGCTCCAGTCAAG 3480
Db 2309 ACAATGCTACTGATGTCAGAACCTTACACAACAGATTCGAAGTTTCTGCTCCAGTCAAG 2368
Qy 3481 GGTGGAATACTACTGAAGGCTCTGTTCTTGTGATCATATA 3518
Db 2369 GGTGGAATACTACTGAAGGCTCTGTTCTTGTGATCATATA 2406

RESULT 6
US-10-600-070-9
; Sequence 9, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitha, Stanislav
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCES: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ. ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ. ID NO 9
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-9

Query Match 46.3%; Score 1720.8; DB 19; Length 2406;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;
Qy 481 ATGGAAGCTCTGATGATCGATCGGCAATTTGCTCTCCCATTCGAAATTAAGCGATTAACA 540
Db 1 ATGGAAGCTCTGATGATCGGCAATTTGCTCTCCCATTCGAAATTAAGCGATTAACA 60
Qy 541 CCGGCGACGACAAAGCTCCGACCTAGCACAACACCTCTACAACTATCTGCTCCGACG 600
Db 61 CCGGCGACGACAAAGCTCCGACCTAGCACAACACCTCTACAACTATCTGCTCCGACG 120
Qy 601 AATGGGCGACCGCTCTCTCTCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
Db 121 AATGGGCGACCGCTCTCTCTCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180
Qy 661 TTGCGACGCGCACACACCGCACCTCGTCTCTGCGACCACTATTAATGATCGCC 720
Db 181 TTGCGACGCGCACACACCGCACCTCGTCTCTGCGACCACTATTAATGATCGCC 240
Qy 721 GAACGCAAGTCCCATCCCATTTGATTTCTACAGATTAATGAGCTCAACACATTTG 780
Db 241 GAACGCAAGTCCCATCCCATTTGATTTCTACAGATTAATGAGCTCAACACATTTG 300
Qy 781 TTAAACGATGGAATCAAGAGCATTCGAAGCTTAGGGTTTCAACCCGCGCAATTCGT 840
Db 301 TTAAACGATGGAATCAAGAGCATTCGAAGCTTAGGGTTTCAACCCGCGCAATTCGT 360
Qy 841 TTGAGGAGGAGCGCTTAAATCAGCGGAGACAGATTTCAAGCTTGGCGAACTCTG 900
Db 361 TTGAGGAGGAGCGCTTAAATCAGCGGAGACAGATTTCAAGCTTGGCGAACTCTG 420
Qy 901 TCTAATCTCGGCTCTGAAAGAGTCAATGAAGTCTTCTTGATGATGAAGAGCTACA 960
Db 421 TCTAATCTCGGCTCTGAAAGAGTCAATGAAGTCTTCTTGATGATGAAGAGCTACA 480
Qy 961 GTGATCACTGATGTTCTTGGGATTAAGTAATTCGATTCGGAATTAAGTTCTTCT 1020
Db 481 GTGATCACTGATGTTCTTGGGAT----- 504
Qy 1021 GTTTTAATTTCAATGATGATTAAGAAAGAACTTTATCTAATGATGAGTTCTGGGGC 1080
Db 505 ----- 518
Qy 1081 TCTCTGATTTGCAAGAGGTTGATGAGCTGAGATAGTTCTTGGGTTGATGAGCTCT 1140
Db 519 TCTCTGATTTGCAAGAGGTTGATGAGCTGAGATAGTTCTTGGGTTGATGAGCTCT 578
Qy 1141 GCTTAAGAGAGGTTGCTTAAGTGTGTTTAAGCAAGATGTGTTTATGATAGGCGCTTCG 1200

Db 579 GCTTAAGAGAGGTTGCTAAGTGTGTTTAAGCAAGATGTGGTTTATGTATAGGCGCTGC 638
 Qy 1201 GTTTCCTGATGCTCGAGGGATGCTATGGCATTTGATCCACTGATTTTATTAATCTGGTTA 1260
 Db 639 GTTTCCTGATGCTCGAGGGATGCTATGGCATTTGATCCACTGATTTTATTAATCTGGTTA 698
 Qy 1261 TGAAGTTGTTGAGGAGCTTTGAGCTTTTACAGATGATTGACCTTGCTTTGGTAATTTG 1320
 Db 699 TGAAGTTGTTGAGGAGCTTTGAGCTTTTACAGATGATTGACCTTTGCTTTGTAATTTG 730
 Qy 1321 ACGAGCGTTGGCTTTATAGAACTTTCTTGATTTGATACCTTTGTTATGAGCTTGCTGA 1380
 Db 731 -----A 731
 Qy 1381 GAGAGAGAGAGCAAGTACCTTGACCGGATTTACGTGCACAAATTGATGACCTTTGGA 1440
 Db 732 GAGAGAGAGAGCAAGTACCTTGACCGGATTTACGTGCACAAATTGATGACCTTTGGA 791
 Qy 1441 AGAGATGACCTCCGCGTTATGCTTGAGCTACCTGGCTTACCGCTGGTGAATGATTA 1500
 Db 792 AGAGATGACCTCCGCGTTATGCTTGAGCTACCTGGCTTACCGCTGGTGAATGATTA 851
 Qy 1501 TGCGAAAAAGCTAATGCTTTAAGCGGTGTCGGAATATTTGCTGCTGTTGAGAGAG 1560
 Db 852 TGCGAAAAAGCTAATGCTTTAAGCGGTGTCGGAATATTTGCTGCTGTTGAGAGAG 911
 Qy 1561 TGAGAGATCAGCTCTGTTGGGGGTTTGACCGGTGAGAAAGTTATGAAATGAGCGGTTTT 1620
 Db 912 TGAGAGATCAGCTCTGTTGGGGGTTTGACCGGTGAGAAAGTTATGAAATGAGCGGTTTT 971
 Qy 1621 ACGAATGACGCTGCTGAGAGATACAGTTTGAATACCTTTTAAATTTCTTTAGCA 1680
 Db 972 ATGATATGACGCTGCTGAGC----- 991
 Qy 1681 TGAATATGCTTAAAGTTTCTCATTTTATATGATGTGTGTGGTGTGAATCTTTTGTGA 1740
 Db 992 -----AGGTTGATCTTTTGTGA 1008
 Qy 1741 GCTAACCCAGCAATATTCAGACAGATCATTTGAAGTTTACGAAGTTGACCTTGCTCTT 1800
 Db 1009 GCTAACCCAGCAATATTCAGACAGATCATTTGAAGTTTACGAAGTTGACCTTGCTCTT 1068
 Qy 1801 GTGCTCAAGCTTTATTTGTTAAGAGCCACCTTTTACAGATGCTGATTAACCAATTC 1860
 Db 1069 GTGCTCAAGCTTTATTTGTTAAGAGCCACCTTTTACAGATGCTGATTAACCAATTC 1128
 Qy 1861 CAGCAACTTCAGACAGCTAAGATGATGCTATGAGATTCCTGCGATGTTGTATGATACA 1920
 Db 1129 CAGCAACTTCAGACAGCTAAGATGATGCTATGAGATTCCTGCGATGTTGTATGATACA 1188
 Qy 1921 CGGAATATTTGAGAGATAGACTTCGGTCTAGAAAGGGGACTCTGTGCACTGCTTATAGGC 1980
 Db 1189 CGGAATATTTGAGAGATAGACTTCGGTCTAGAAAGGGGACTCTGTGCACTGCTTATAGGC 1248
 Qy 1981 AAAGTTATGAAATGCTGATGCTGTTGGGCTTAAACATGAGGATTTCAATATATAGGAAT 2040
 Db 1249 AAAGTTATGAAATGCTGATGCTGTTGGGCTTAAACATGAGGATTTCAATATATAGGAAT 1308
 Qy 2041 CCAAGCTATTTGAGAGTTGTTTGAAGAAATCAATGCTGATGCAATATGATGCTCCCT 2100
 Db 1309 CCAAGCTATTTGAGAGTTGTTTGAAGAAATCAATGCTGATGCAATATGATGCTCCCT 1368
 Qy 2101 GGAATATGCAATTTGTTGAGAACTGTGTGCAAGGGGTTGTCTTTCTTCAAGGTTCAAGAC 2160
 Db 1369 GGAATATGCAATTTGTTGAGAACTGTGTGCAAGGGGTTGTCTTTCTTCAAGGTTCAAGAC 1428
 Qy 2161 ACCCAAAATATAAAAATTTAAATCTGGGGACATATATGATGATCTTATGCTTTTGAATAC 2220
 Db 1429 ACCCAAAATATAAAAATTTAAATCTGGGGACATATATGATGATCTTATGCTTTTGAATAC 1488
 Qy 2221 TTGGAAGAAGTGAAGTATGCTCAGGGTTCTCTTTAGCTGCTGTCGACATATGCGAAG 2280

Db 1489 TTGGAAGAAGTGAAGTATGCTAAGGTTTCTCTTACGCTGCTGACATATGCGAAG 1548
 Qy 2281 ATTGAGCCGAGCATGAGAAAGTATGCTATGACAGGCACTGCGAAGATTTTCTTCC 2340
 Db 1549 ATTGAGCCGAGCATGAGAAAGTATGCTATGACAGGCACTGCGAAGATTTTCTTCC 1608
 Qy 2341 CGCTATACAGATAGAAATCTGGCTGAACCCAGAGATGCGAAGAGACAGTGTATAGTGA 2400
 Db 1609 CGCTATACAGATAGAAATCTGGCTGAACCCAGAGATGCGAAGAGACAGTGTATAGTGA 1668
 Qy 2401 GATCTGTTGGTAAACAATGATAGGCGCTGATATGATGAGCTGTGTCTTTATTCGAAGCT 2460
 Db 1669 GATCTGTTGGTAAACAATGATAGGCGCTGATATGATGAGCTGTGTCTTTATTCGAAGCT 1728
 Qy 2461 GTTAAAGCCCTGTGAAACTTTGAACTTAATGCAATTGAGCGTGGGGGTCTCAGAG 2520
 Db 1729 GTTAAAGCCCTGTGAAACTTTGAACTTAATGCAATTGAGCGTGGGGGTCTCAGAG 1788
 Qy 2521 AGTAGCGTTGATGAAATCACTAGTTGAATGCTGCTGCTGATATGTTAAAGAGGCAAGT 2580
 Db 1789 AGTAGCGTTGATGAAATCACTAGTTGAATGCTGCTGCTGATATGTTAAAGAGGCAAGT 1848
 Qy 2581 GTGAAGATCCTAGCTGCTGTGTGTGCAATTTGACTGATTTCACTGTTCAAGCCAGAGTAT 2640
 Db 1849 GTGAAGATCCTAGCTGCTGTGTGTGCAATTTGACTGATTTCACTGTTCAAGCCAGAGTAT 1908
 Qy 2641 TTTCTTAAAAAGCACTCATCTTTTCAACGCAAGATATGCTTTCTTATGCAATTCGAT 2700
 Db 1909 TTTCTTAAAAAGCACTCATCTTTTCAACGCAAGATATGCTTTCTTATGCAATTCGAT 1968
 Qy 2701 GTGCTACCAATGATATGATTAATGATGCAATTTTCAATATCTGATGCTCAATAATA 2760
 Db 1969 GTGCTACCA----- 1978
 Qy 2761 TGCTGTTTGTGAGCTAAGAACATAGTCCACTTAATACATGCTCCAAAGTTGTACC 2820
 Db 1979 ----- 1978
 Qy 2821 AAGATTAAACAAGTTGCTGATGAATTTTCACTAATATATGCTGCTGAATTTTGTGATA 2880
 Db 1979 ----- 1978
 Qy 2881 CTGTAGACAGAAATGTAATTTTCACTCTCAACATTTCTGTTAGATTAACGTAGAGATTAG 2940
 Db 1979 ----- 1978
 Qy 2941 AGATTCCTTAGTGTGCTTGTCCAACTTTTCTTCTTGATTTTCTTTTGCATTT 3000
 Db 1979 -----T 1979
 Qy 3001 AGGTCATGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATATGCTAGACTGACGA 3060
 Db 1980 AGGTCATGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATATGCTAGACTGACGA 2039
 Qy 3061 GAATATATGATCCAGAGTGGCAGAAATTAAGTCTGCTGCTTTTGGGCTGATCACCCGAT 3120
 Db 2040 GAATATATGATCCAGAGTGGCAGAAATTAAGTCTGCTGCTTTTGGGCTGATCACCCGAT 2039
 Qy 3121 AGAAATGTTACAGAGGTGAGGGAATTAATCTAACATTCATCAATGTGTGAAGACTGT 3180
 Db 2100 AGAAATGTTACAG----- 2113
 Qy 3181 TGAACATGATTAATAGTGTGCTGCTTGTGATTTCTATTAATATGATTTTGAATGGGC 3240
 Db 2114 -----AGTTTTGAATGGGC 2128
 Qy 3241 GAATGCTGAAGATTTGAGCTGACAGAGCACTGAAACTGCGACGCTTGGGTTGATG 3300
 Db 2129 GAATGCTGAAGATTTGAGCTGACAGAGCACTGAAACTGCGACGCTTGGGTTGATG 2188
 Qy 3301 ATTATACACTGTTGAACATATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCCGTG 3360
 Db 2189 ATTATACACTGTTGAACATATCTGTTGACAGTGTGACAGTCTCAGCAGATGGAACCCGTG 2248

QY 3361 CTCTGGGAAAGCACTCTGAGAGAGCTGCTGTCTATCTGATGATTTTGGTTCATCCAGAAA 3420
| | | | |
Db 2249 CTCTGGGAAAGCACTCTGAGAGAGCTGCTGTCTATCTGATGATTTTGGTTCATCCAGAAA 2308
| | | | |
QY 3421 ACAATGCTACTGATGTGCAGAACTCAACAACAAGATACGAAGTTTCTGTGCTCAAGTCAAG 3480
| | | | |
Db 2309 ACAATGCTACTGATGTGCAGAACTCAACAACAAGATACGAAGTTTCTGTGCTCAAGTCAAG 2368
| | | | |
QY 3481 GGTTGAAAAATCACTGAAGGCTCTGTCTTCTGATCATATA 3518
| | | | |
Db 2369 GGTTGAAAAATCACTGAAGGCTCTGTCTTCTGATCATATA 2406
| | | | |
RESULT 7
US-10-600-070-128
/ Sequence 128, Application US/10600070
/ Publication No. US20040139500A1
/ GENERAL INFORMATION:
/ APPLICANT: Oseeryoung, Katherine W.
/ APPLICANT: Vilha, Stanislav
/ APPLICANT: Kokharova, Olga A.
/ APPLICANT: Geo, Hongso
/ TITLE OF INVENTION: Plaeid Division and Related Genes and Proteins, and Methods of
/ FILE REFERENCE: MSU-08153
/ CURRENT APPLICATION NUMBER: US/10/600,070
/ CURRENT FILING DATE: 2003-06-20
/ NUMBER OF SEQ ID NOS: 206
/ SOFTWARE: Patentin version 3.2
/ SEQ ID NO 128
/ LENGTH: 2406
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
US-10-600-070-128
Query Match 46.8%; Score 1717.6; DB 19; Length 2406;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 632; Gaps 5;
QY 481 ATGGAAGCTTGAAGTCAAGTCCGAGTGGTGTCTCTCCCATTTCCAAATATGCGGATTAACA 540
| | | | |
Db 1 ATGGAAGCTTGAAGTCAAGTCCGAGTGGTGTCTCTCCCATTTCCAAATATGCGGATTAACA 60
| | | | |
QY 541 CCGGCGAGCAAGAGCTCCGAGCGTAGGCACAACCTCTCAACTATCTGCTCGGCGAGC 600
| | | | |
Db 61 CCGGCGAGCAAGAGCTCCGAGCGTAGGCACAACCTCTCAACTATCTGCTCGGCGAGC 120
| | | | |
QY 601 AAATGGCCGACGCTTCTCTCGACTTCATTTCACTTCGATTCCTCTCTCTCTCC 660
| | | | |
Db 121 AAATGGCCGACGCTTCTCTCGACTTCATTTCACTTCGATTCCTCTCTCTCTCTCC 180
| | | | |
QY 661 TTGGCCACCGCCACACACACCGGCACTCTGTCTCTTGCCACATTTATTTGATGCTGCC 720
| | | | |
Db 181 TTGGCCACCGCCACACACACCGGCACTCTGTCTCTTGCCACATTTATTTGATGCTGCC 240
| | | | |
QY 721 GAAGCGCAGCTCCCATCCCATTTGATTTCTACCAAGGATTTAGAGGCTCAAAACATTTTC 780
| | | | |
Db 241 GAAGCGCAGCTCCCATCCCATTTGATTTCTACCAAGGATTTAGAGGCTCAAAACATTTTC 300
| | | | |
QY 781 TTAAACGATGAATTCAGAAAGCAATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGAT 840
| | | | |
Db 301 TTAAACGATGAATTCAGAAAGCAATTCGAAGCTAGGGTTTCGAAACCGCGCAATTCGAT 360
| | | | |
QY 841 TTCAAGGACGACGCTTTATTCAGCCGAGACATTTCTTCAAGTGTCTTGCGAAACTCTG 900
| | | | |
Db 361 TTCAAGGACGACGCTTTATTCAGCCGAGACATTTCTTCAAGTGTCTTGCGAAACTCTG 420
| | | | |
QY 901 TCTAATCTCGGTCTAGAGAGAGTACATGAAGGCTCTTCTTGATGATGAAGAGCTACA 960
| | | | |
Db 421 TCTAATCTCGGTCTAGAGAGAGTACATGAAGGCTCTTCTTGATGATGAAGAGCTACA 480
| | | | |
QY 961 GTGATCACTGATGTTCTTGGGATTAAGTATTCGATTCGAAATATTAAGTTTCTTC 1020
| | | | |

Db 481 GTGATCACTGATGTTCTTGGGAT----- 504
| | | | |
QY 1021 GTTTTAATTTCAAGAAATGGAATTAAGAAAGCACTTTATCTAGTGAAGGTTCTGGGGC 1080
| | | | |
Db 505 -----AAGGTTCTGTGTC 518
| | | | |
QY 1081 TCTCTGTGATTTGCAAGAAAGGTGTGAGACTGATAGTTCCTTGGGTTGGTGAAGGCTCT 1140
| | | | |
Db 519 TCTCTGTGATTTGCAAGAAAGGTGTGAGACTGATAGTTCCTTGGGTTGGTGAAGGCTCT 578
| | | | |
QY 1141 GCTTAAAGAGAGGTTGCTTAAGTCTTTTAAAGCAAGATGTGTTTAACTTATGCGCTTGC 1200
| | | | |
Db 579 GCTTAAAGAGAGGTTGCTTAAGTCTTTTAAAGCAAGATGTGTTTAACTTATGCGCTTGC 638
| | | | |
QY 1201 GTTTCGATGCTCGAGGGAGTGTATGAGCTTGATCCACTGATTTATTACTGGTTA 1260
| | | | |
Db 639 GTTTCGATGCTCGAGGGAGTGTATGAGCTTGATCCACTGATTTATTACTGGTTA 698
| | | | |
QY 1261 TGAGTTTGTGAGAGCTTTGAAGCTTTTACAGATGATTTGACTTGCTTGGTAAATTTG 1320
| | | | |
Db 699 TGAGTTTGTGAGAGCTTTGAAGCTTTTAC----- 730
| | | | |
QY 1321 ACGAGCGTTGGCTTTATTAAGAACTTTCTGATTTGATCTTGTATGAGTCTGTGCTA 1380
| | | | |
Db 731 -----A 731
| | | | |
QY 1381 GGAGGAAGAGCAAGTAGCTTGACCGGATTTACGTGCAAAATGATGAGACTTTGGA 1440
| | | | |
Db 732 GGAGGAAGAGCAAGTAGCTTGACCGGATTTACGTGCAAAATGATGAGACTTTGGA 791
| | | | |
QY 1441 AGAGATCACTCCGCGTATGCTCTGAGACTCTTGCGCTTACCGCTTGATGATTAACGC 1500
| | | | |
Db 792 AGAGATCACTCCGCGTATGCTCTGAGACTCTTGCGCTTACCGCTTGATGATTAACGC 851
| | | | |
QY 1501 TGGAAAAAGCTAAATGTTTAAAGCGGTGCGGAAATTTTGTGCTCTTGAGAGAG 1560
| | | | |
Db 852 TGGAAAAAGCTAAATGTTTAAAGCGGTGCGGAAATTTTGTGCTCTTGAGAGAG 911
| | | | |
QY 1561 TGGAGATCACTCTTGTGCGGGTTTGAACCGGTGAGAAATTTATGATGAGGCGTTTTC 1620
| | | | |
Db 912 TGGAGATCACTCTTGTGCGGGTTTGAACCGGTGAGAAATTTATGATGAGGCGTTTTC 971
| | | | |
QY 1621 ACGAATGACAGCTGCTGAGCAGGATACAGTTAGATCACTTTTAAATTTCTTTAGCA 1680
| | | | |
Db 972 ACGAATGACAGCTGCTGAGC----- 991
| | | | |
QY 1681 TGATATTAATTTAGGTTTCTCATTTTAATGATGTTGTGATGATCTTTTGTGA 1740
| | | | |
Db 992 -----AAGTTGATCTTTTGTGA 1008
| | | | |
QY 1741 GTTACCCCAAGCAATTTCCAGAGAGTCAATTTGAAGTTTACGAATGTCATTTGCTCTT 1800
| | | | |
Db 1009 GTTACCCCAAGCAATTTCCAGAGAGTCAATTTGAAGTTTACGAATGTCATTTGCTCTT 1068
| | | | |
QY 1801 GTGGCTCAAGCTTTTATTTGTTAAAGCAACACCTTTTACAGATGCTGTAAACAAATTC 1860
| | | | |
Db 1069 GTGGCTCAAGCTTTTATTTGTTAAAGCAACACCTTTTACAGATGCTGTAAACAAATTC 1128
| | | | |
QY 1861 CAGCACTTCAGAGGCTAAGTAAATGCTAATGAGATTCCTGCGATGTTGATGATGATACA 1920
| | | | |
Db 1129 CAGCACTTCAGAGGCTAAGTAAATGCTAATGAGATTCCTGCGATGTTGATGATGATACA 1188
| | | | |
QY 1921 CCGAATTAATTTGGAGATTAAGCTTCGCTTGAAGAAAGGGAATCTGTGCACTGCTTAATAGC 1980
| | | | |
Db 1189 CCGAATTAATTTGGAGATTAAGCTTCGCTTGAAGAAAGGGAATCTGTGCACTGCTTAATAGC 1248
| | | | |
QY 1981 AAGTTGATGAATGCCGATATGCTTGGGCTTGAAGAGAGATTACAAATATAGGAAT 2040
| | | | |
Db 1249 AAGTTGATGAATGCCGATATGCTTGGGCTTGAAGAGAGATTACAAATATAGGAAT 1308
| | | | |
QY 2041 CCAAGCTTATTTGAGAGTTTGTTTTGAAGAAATTCAAATCGTATGACAAATGATGATCTCCCT 2100
| | | | |

Db 1309 CCAGTATTTGTGAGTTGTTTGGAGAAATTCAAATCGATGACATGATGATCTCCCT 1368
 QY 2101 GGCATATGCAAAATTTGTGGAAACCTGGTTGGCAGGGGTTTCTTTCTAGTTCAGAGAC 2160
 Db 1369 GGCATATGCAAAATTTGTGGAAACCTGGTTGGCAGGGGTTTCTTTCTAGTTCAGAGAC 1428
 QY 2161 ACCAAGAATAAAATTTAACTCGGGGACTACTATGATGATCCTATGTTTGAATTAAC 2220
 Db 1429 ACCAAGAATAAAATTTAACTCGGGGACTACTATGATGATCCTATGTTTGAATTAAC 1488
 QY 2221 TTGGAAGAGTGAAGGTAGTTCAGGGTTCCTTTAGCTGCTGCTGCAACTATGAGCAGG 2280
 Db 1489 TTGGAAGAGTGAAGGTAGTTCAGGGTTCCTTTAGCTGCTGCTGCAACTATGAGCAGG 1548
 QY 2281 ATTGAGGCCAGCATGAGGATGAGTATGAGGCACTGCGAGAAAGTTTTCCTTCC 2340
 Db 1549 ATTGAGGCCAGCATGAGGATGAGTATGAGGCACTGCGAGAAAGTTTTCCTTCC 1608
 QY 2341 CGCTATACAGATGAAACTCGGCTGAACCAAGGATGCAAGAGACAGTGTATGATGA 2400
 Db 1609 CGCTATACAGATGAAACTCGGCTGAACCAAGGATGCAAGAGACAGTGTATGATGA 1668
 QY 2401 GATCTGTTGGTAAACAATGATGAGCCGTGATGATGAGCCGTGATGATGAGGAACT 2460
 Db 1669 GATCTGTTGGTAAACAATGATGAGCCGTGATGATGAGCCGTGATGATGAGGAACT 1728
 QY 2461 GTAAAGACCTCTGAAAACTTTGAAACTAATGATTAATGCAATTCAGAGTGGGGTCTCAG 2520
 Db 1729 GTAAAGACCTCTGAAAACTTTGAAACTAATGATTAATGCAATTCAGAGTGGGGTCTCAG 1788
 QY 2521 AGTAGCGTTGATGAACTACTGTTGAAATGCTGTTGATGATGATGATGATGATGATGAT 2580
 Db 1789 AGTAGCGTTGATGAACTACTGTTGAAATGCTGTTGATGATGATGATGATGATGATGAT 1848
 QY 2581 GTGAAGATCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
 Db 1849 GTGAAGATCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1908
 QY 2641 TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGATGATGATGAT 2700
 Db 1909 TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGATGATGATGAT 1968
 QY 2701 GTGCGTACCAAGTATGATTAATGATGCAATTTGATATGATGATGATGATGATGATGAT 2760
 Db 1969 GTGCGTACCAAGTATGATTAATGATGCAATTTGATATGATGATGATGATGATGATGAT 1978
 QY 2761 TGCTTGTGTTGTGAGTGAAGACATAGTCCCACTTAATACATGTCCTCAAAAGTTGATCC 2820
 Db 1979 TGCTTGTGTTGTGAGTGAAGACATAGTCCCACTTAATACATGTCCTCAAAAGTTGATCC 1978
 QY 2821 AAGATTAAACAAGTTGCTGAGTAAATTTCACTAATTATGCTGCTGGAATTTTTCATCAA 2880
 Db 1979 TGCTTGTGTTGTGAGTGAAGACATAGTCCCACTTAATACATGTCCTCAAAAGTTGATCC 1978
 QY 2881 CTGTAGACAGAAATGTAATTTCACTCAACATTTGTTTGAATTAAGTAGATTAAG 2940
 Db 1979 TGCTTGTGTTGTGAGTGAAGACATAGTCCCACTTAATACATGTCCTCAAAAGTTGATCC 1978
 QY 2941 AGATTGCTTAGTGTGCTTGTTCACACTTTTCTTCTGATTTTCTTTTTCATGATTT 3000
 Db 1979 TGCTTGTGTTGTGAGTGAAGACATAGTCCCACTTAATACATGTCCTCAAAAGTTGATCC 1979
 QY 3001 AGGCTCAGTCAAGCTGACGATTCAGAGACCTTCCAGAAATGATGATGATGATGATGATGAT 3060
 Db 1980 AGGCTCAGTCAAGCTGACGATTCAGAGACCTTCCAGAAATGATGATGATGATGATGATGAT 2039
 QY 3061 GAATATAGTATCCAGAGTGAAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 3120
 Db 2040 GAATATAGTATCCAGAGTGAAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 2099
 QY 3121 AGAAATTTTACCAAGGTGAGGAAATTAATCTCAATTCATCAATGATGATGATGATGATGAT 3180
 Db 2100 AGAAATTTTACCAAGGTGAGGAAATTAATCTCAATTCATCAATGATGATGATGATGATGAT 2113

QY 3181 TGGACATGATTATGATGCTGAGCCCTGTTGATTCGTTATTTATAGTTTGGATGGC 3240
 Db 2114 -----AGTTTGGATGGC 2128
 QY 3241 GAATGCTGAAGATTGGACATGACAGACAGCTGAACCTGCGCAGCTTGGTGGTTATG 3300
 Db 2129 GAATGCTGAAGATTGGACATGACAGACAGCTGAACCTGCGCAGCTTGGTGGTTATG 2188
 QY 3301 ATTATACATGTTGAACTATCTGTTGACAGTGTGACAGTCTGAGCAGATGGAACCCGTG 3360
 Db 2189 ATTATACATGTTGAACTATCTGTTGACAGTGTGACAGTCTGAGCAGATGGAACCCGTG 2248
 QY 3361 CTCTGGTGAAGCAACTCTGAGAGAGTCTGCTGATCTGATCTGATTTGTTGATTCAGAAA 3420
 Db 2249 CTCTGGTGAAGCAACTCTGAGAGAGTCTGCTGATCTGATCTGATTTGTTGATTCAGAAA 2308
 QY 3421 ACAATGCTACTGATGTGACAACTTACACAAAGATTAAGATTTTGTGTCAGATCAG 3480
 Db 2309 ACAATGCTACTGATGTGACAACTTACACAAAGATTAAGATTTTGTGTCAGATCAG 2368
 QY 3481 GGTGAAAAATCACTGAAGGCTCTGTTCTTGCATCATTA 3518
 Db 2369 GGTGAAAAATCACTGAAGGCTCTGTTCTTGCATCATTA 2406
 RESULT 8
 US-10-600-070-132/c
 ; Sequence 132, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Viltha, Stanislaw
 ; APPLICANT: Koksharova, Olga A.
 ; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; CURRENT FILING DATE: 2003-06-20
 ; NUMBER OF SEQ ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 132
 ; LENGTH: 561
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (127)..(127)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (520)..(520)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (541)..(541)
 ; OTHER INFORMATION: n is a, c, g, or t
 ; US-10-600-070-132
 Query Match 11.0%; Score 402.2; DB 19; Length 561;
 Best Local Similarity 82.2%; Pred. No. 4.6e-95;
 Matches 514; Conservative 0; Mismatches 20; Indels 91; Gaps 1;
 QY 2970 TTTCTTCTTGAATTTTCTTTTCAATTTAGGCTCAGTCAAGCTGACGATTCAGAG 3029
 Db 534 TTTCTTCTTGAATTTTCTTTTCAATTTAGGCTCAGTCAAGCTGACGATTCAGAG 475
 QY 3030 CACTTCCCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3089
 Db 474 CACTTCCCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 415
 QY 3090 AGTCTGAGCTTTTGGCTGATACCGCATAGAAATGTTACAGAGTGAAGGAATTA 3149


```

QY 767 CTCACACATTTCTTAACCGATGAAATCAGAGAGCATTCGAACTGAGGTTTCGAAAC 826
DB 407 CCGAACCGCATCTTCTCGGGGAGCGGATTCGAGGGCTTACGAGGCCAAAGTTCTCGAAGC 466
QY 827 GCGCGCAATTCGGTTTCAGGAGCAGCGTTTAATCAGCCGAGACAGATTCTTCAAGCTG 886
DB 467 CTCCTCAGTACGCGCTTACGACACGCTTTAATCAGCCGCGCAATCTCTCAACAG 526
QY 887 CTTCGAAACCTCTCTAATCTCGGTCTAGAAAGATCAATGAAGTCTTCTGATG 946
DB 527 CTTCGAAACCTCTAATCTCTAATCTCTAATCTCTAATCTCTAATCTCTAATCTCTA 586
QY 947 ATGAAGAAGCTACAGTACATCATGATGTTCTTCGGATTAAGTAATTCGATTCGAAAT 1006
DB 587 ACGAAGAAGGCGCATCTGACCTCAATCCCTT----- 619
QY 1007 AATTAAGTTTCTTCTTTAATTTCAATGATGATTAAGAAAGAACTTTAATCTAGTG 1066
DB 620 -----TCGAC 624
QY 1067 AAGGTTCTGGGGCTCTGTGTATTCGAAAGAGTGTGAGACTGAGATGTTCTTCG 1126
DB 625 AAGGTTCTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 684
QY 1127 GTTGTGAGGCTCTGCTTAAGAGAGGTTGCTTAAGTCTTAAGCAAGATGTGTGTTTA 1186
DB 685 ATTGGGAGGTTTGTCTTAAGAGAGGTTGCTTAAGCAAGATGTGTGTTG 744
QY 1187 GTTATGCGCTTGGGTTCTGTGATGTCTGAGGAGTCTATGCAATGATTCACCTGAT 1246
DB 745 GCTATGCGCATCTGCAATTTGTGATGCTCAAGGAGATCTATGCTTGTCTCCACCGAT 804
QY 1247 TTTATTAATGTTATGATGTTTGTGAGAGGTTTGAAGCTTTTACAGTAATGTTGACT 1306
DB 805 TTCAATGCGGCTGTGTGATGCTGTGAGAGGCTGTGAAGCTTTTGAAGAA----- 855
QY 1307 GCTTTGGTAAATTTGACGAGCGTTGCTTTAATGAACCTTTCTGATTTGATTTGTTA 1366
DB 856 ----- 855
QY 1367 TTGAGTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1426
DB 856 -----GAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
QY 1427 GATGAGCTTTGAGAGAGATCATCTCCGCTTATGCTTTGAGAGTACTTGGCTTACCGCT 1486
DB 898 GATGAGAGCGCTGAGAGAGATTAACCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTG 957
QY 1487 GGTGATGATTAACGCTGCGAAAGAACTAATGTGTTAAGCGGTGTGCGGAATATTTGTG 1546
DB 958 GATGAGAGAGATGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
QY 1547 TCTGTGAGAGAGAGTGAAGCATCATCTTGTGAGGAGTTTGAACCGCTGAGAGATTTAG 1606
DB 1018 GCGGTTGTGTGAGAGGAGGTGAGAGCAATGTCGCGGAGTTTACCGGTGAAGATTTATG 1077
QY 1607 AATGAGAGCGTTTATTCAGATGAGAGCTGTGTGAGAGAGATTAACGTTTATGATACC 1660
DB 1078 AATGAGAGCATTTTACATGATGAGAGCGGCTGAGAGAGTTGAATTTTGTAGCC 1131

```

RESULT 11
US-10-600-070-135

```

; Sequence 135, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153

```

```

; CURRENT APPLICATION NUMBER: US/10/600.070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 135
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Medicago truncatula
US-10-600-070-135

```

```

Query Match 5.7%; Score 208.8; DB 19; Length 660;
Best Local Similarity 60.3%; Pred. No. 5.8e-44;
Matches 429; Conservative 0; Mismatches 207; Indels 76; Gaps 2;

```

```

QY 561 ACGTAGCCACAAACCTCTCAATCTATCTGTGCGCAGCAAAATGGGCGGACGCTTCT 620
DB 17 ACCTAACCGTCTCATCTCTCGCGCTGTGCGCACAGTAATGGGCGGACGCTCAT 76
QY 621 CTCGACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 680
DB 77 TTCCGATTTCCAAATTCCTCGGCGACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136
QY 681 CGCCACTCTGTCTCTCTGCAACCATTAATGATGTGTCCGAGCGCACGTCCTCCATCC 740
DB 137 AGTCACTCTC-----ACTCTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190
QY 741 CATGATTTTACAGGATTTAGAGGCTCAAGCAATTTCTTAACGATGATGATGATGATG 800
DB 191 TCTGACCTGTACAAATTCCTCGGCGCAAGAGCAATTTCTCGGTATGATTTGAGAG 250
QY 801 AGCATTCGAAGCTTGAAGGTTTGAAGAGCGGCAATTCGTTTACGAGAGAGAGCTTTAAT 860
DB 251 AGCTTATGAAGCGAAATTCGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 310
QY 861 CAGCGGAGAGAGATTTCTTAAGTGTCTGTGAGAACTCTGTCTAATCTCTGCTCTAGAG 920
DB 311 TAGCTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 370
QY 921 AGAGTACATGAAGGCTTCTGTATGATGATGATGATGATGATGATGATGATGATGATG 980
DB 371 AGAGTATATCAAAAGCTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 405
QY 981 GATTAAGTAAATTCGATTCGAAATTAAGTTTCTGTGTTAATTTATGATGATGATG 1040
DB 406 -----MAGTAGAGAACTT 420
QY 1041 ATTAAGAGAGAACTTTATCTAGTGAAGGTTCTGTGAGGCTCTGTGTATGCAAGAG 1100
DB 421 CAATTCATCTGAATCCCTTTGCAAAAGTTCTGTGAGGCTGTGTGAGGCTGTGTGCAAG 480
QY 1101 GTGTGAGTGAATGTTCTTCTGAGTGTGTGAGGCTGTGTGAGGCTGTGTGAGGCTGT 1160
DB 481 CTGAGAGAGAGAGGTTGTCTTCTGATGATGAGGAGGTTTACATGAGAGAGAGTTACGA 540
QY 1161 AGTGTGTTAAGCAAGATGTGTTTATGTTATGAGGCTGTGTGAGGCTGTGTGAGGCTGT 1220
DB 541 AGATGTTTAAAGCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 1221 ATGCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1272
DB 601 ATGCTATGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 652

```

RESULT 12
US-10-600-070-126

```

; Sequence 126, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of

```



```
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen L.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51770)B
; CURRENT APPLICATION NUMBER: US/09/732,627A
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4930
; SEQ ID NO 2154
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(439)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3493-032-PI-M1-F8
; US-09-732-627A-2154

Query Match          4.8%; Score 175; DB 11; Length 439;
Best Local Similarity 65.0%; Pred. No. 3.7e-35;
Matches 278; Conservative 0; Mismatches 141; Indels 9; Gaps 1;

QY 499 GTCCGCAATGGTCTCTCCCATTCATTAATGCGATTACCAACCGGCGAGCAAAAGCTC 558
   |||||
DB 12 GTTAGCATGGGCTGTGACATCCGACACTTGTCTTATTGACACCTGTAAACCTCTCC 71
QY 559 CGACGTAGCCACAACCTCTTACACTATCTGTCCGCGCAAAATGGGCGAGCGCTCTT 618
   |||||
DB 72 AACTCCACCAACCGTTCAACCTGCTGTCTCCGCACTAAATGGGCGAGCGCTCTC 131
QY 619 CTCCTCCGACTTCAATTTCACTCCG-----ATTCTCTCTCTCTCTCTCTCTGCGCAC 669
   |||||
DB 132 CTGCTGATCTTCATTTCTCTCCGCGCGCGATTAATTCGTTCTCTCTCTCTCTCA 191
QY 670 GCCACACCAACCGGCACTCTGTCTCTGCGCACATTAATTAATTCGCGAAAGCGCAC 729
   |||||
DB 192 ACCGCGCACTTTTCTCCCGCTTACCTCCCTTCTCTGCGCGCTCTCTCTCGGAAAGCGCAC 251
QY 730 GTCCCATCCCATTTGATTTCTTACAGGTAATGAGAGCTCAACACATTTCTTAACGAT 789
   |||||
DB 252 GTTTCATTTCCCTTGAATTTCTTACAGGTTTGAAGCGAGACATTTCTTAGTAT 311
QY 790 GGATCGAAGAGATTCGAGCTAGGTTTCGAAACCGCGCAATTCGTTTCAGCGAC 849
   |||||
DB 312 GGAATTGAAGAGCTTGAAGCAAGGTTTCGAAACCGCTCAATATGAGTTCAAGTCAA 371
QY 850 GACGCTTAATCAACCGGAGAGAGATTTCTCAAGCTGTTGCGAAACTCTGTAAATCT 909
   |||||
DB 372 GACACCATTAATGAGCGAAGACAGATTTCTTAGCTGCTGTGAAGCCCTATTAACCT 431
QY 910 CGGCTTAG 917
   |||||
DB 432 GGCTCTAG 439
   |||||

RESULT 15
US-10-600-070-175
; Sequence 175, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oosteryoung, Katherine W.
; APPLICANT: Vitsha, Stanislav
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: NSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 175
```

```
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Gossypium arboreum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (528)..(528)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (536)..(536)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (540)..(540)
; OTHER INFORMATION: n is a, c, g, or t
; US-10-600-070-175

Query Match          4.4%; Score 161; DB 19; Length 545;
Best Local Similarity 69.6%; Pred. No. 2.1e-31;
Matches 218; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 3224 ATAGCTTTGATGGCGAATGCTGAAGATTTGACATGACAGACAGCTGAAACTGCCCA 3283
   |||||
DB 82 AGAGTTCTGATGATGATCAATGTTGAAGACATGACAGATGTCAGCGCAATGCTCA 141
QY 3284 GCTTGGGTTGTTATGATTATACACTGTTGAACATACTGTGACAGTGAAGTCTC 3343
   |||||
DB 142 GCTTGGTGGTATATGATATGATCTACTGAACTGCAATGCAATGCAATGCTTTC 201
QY 3344 AGCAGATGAACCCGCTCTCTGTGGAAGCAACTGTGAGAGTCTGCTGTATCTGA 3403
   |||||
DB 202 ACTGATGAGCGAGGAGACTGTGATGAAAGTACTGTGAAGAAATCCACTGCTGACTGA 261
QY 3404 TTGCTTCATTCAGAAAACAAATGCTACTGATGTCAAGAACTTACACAAACAGATACGAAT 3463
   |||||
DB 262 TGTTTCATTCATTCGAGAAACAAATGCTCTTAATGTAACTCTTACACCAACGATATGAGAT 321
QY 3464 TTTCGTGTCAGAGTCAAGGTGGAATATCACTGAAGGCTCTGTGATCATATATATAC 3523
   |||||
DB 322 GTTGTGTTCAACTCAGGCTGGAATATCACTGAAGGATGTGTCAAAATCTTAACTATG 381
QY 3524 TCATATGTAGCAT 3536
   |||||
DB 382 ATGTATTAAGCAT 394
   |||||
```

Search completed: June 10, 2005, 01:37:37
Job time : 2087 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2005, 15:25:22 ; Search time 575 Seconds
(without alignments)
10435.177 Million cell updates/sec

Title: US-10-600-070-3

Perfect score: 3667
Sequence: 1 tgcctcgcatgaaggaat.....ctataacataagggctacaa 3667

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.6	1.6	7218	1 US-08-232-463-14	Sequence 14, Appl
2	54	1.5	1141	4 US-09-806-708B-22	Sequence 22, Appl
3	50.2	1.4	57280	4 US-09-949-016-11796	Sequence 11796, A
4	50.2	1.4	57280	4 US-09-949-016-12843	Sequence 12843, A
5	50.2	1.4	57280	4 US-09-949-016-12844	Sequence 12844, A
6	50.2	1.4	57280	4 US-09-949-016-12846	Sequence 12846, A
7	50.2	1.4	57280	4 US-09-949-016-13542	Sequence 13542, A
8	50.2	1.4	57280	4 US-09-949-016-13543	Sequence 13543, A
9	50.2	1.4	57280	4 US-09-949-016-13544	Sequence 13544, A
10	50.2	1.4	57280	4 US-09-949-016-13545	Sequence 13545, A
11	50.2	1.4	57280	4 US-09-949-016-14633	Sequence 14633, A
12	50.2	1.4	57280	4 US-09-949-016-14634	Sequence 14634, A
13	50.2	1.4	57280	4 US-09-949-016-14635	Sequence 14635, A
14	50.2	1.4	57280	4 US-09-949-016-14636	Sequence 14636, A
15	50.2	1.4	57280	4 US-09-949-016-14637	Sequence 14637, A
16	50.2	1.4	57280	4 US-09-949-016-14638	Sequence 14638, A
17	50.2	1.4	57280	4 US-09-949-016-14639	Sequence 14639, A
18	50.2	1.4	57280	4 US-09-949-016-14640	Sequence 14640, A
19	49.4	1.3	913	1 US-08-217-327-3	Sequence 3, Appl
20	49.4	1.3	913	1 US-08-885-970A-3	Sequence 3, Appl
21	49.4	1.3	913	1 US-08-298-687A-3	Sequence 3, Appl
22	49.4	1.3	913	1 US-08-530-787-2	Sequence 2, Appl
23	49.4	1.3	913	1 US-08-238-829-3	Sequence 3, Appl
24	49.4	1.3	913	2 US-08-787-335-2	Sequence 2, Appl
25	49.4	1.3	1984	1 US-07-885-970A-25	Sequence 25, Appl
26	49.4	1.3	1985	1 US-08-298-687A-25	Sequence 25, Appl
27	49.4	1.3	1985	1 US-08-232-463-25	Sequence 25, Appl

28	49.2	1.3	324	4 US-09-547-693-234	Sequence 234, Appl
C 29	48.4	1.3	152132	4 US-09-949-016-13845	Sequence 13845, A
C 30	48.4	1.3	152145	4 US-09-949-016-12371	Sequence 12371, A
C 31	47.4	1.3	1141	4 US-09-806-708B-22	Sequence 22, Appl
C 32	46.8	1.3	114793	4 US-10-148-806-3	Sequence 3, Appl
C 33	45	1.2	49818	4 US-09-949-016-12458	Sequence 12458, A
C 34	45	1.2	49829	4 US-09-949-016-14011	Sequence 14011, A
C 35	44.2	1.2	4403765	3 US-09-103-840A-2	Sequence 2, Appl
C 36	44.2	1.2	4411529	3 US-09-103-840A-1	Sequence 1, Appl
C 37	43.8	1.2	832	4 US-09-621-976-2813	Sequence 2813, Ap
C 38	43.8	1.2	832	4 US-09-621-976-2813	Sequence 2813, Ap
C 39	42.8	1.2	1055	4 US-09-806-708B-23	Sequence 23, Appl
C 40	42.8	1.2	60990	4 US-09-949-016-14080	Sequence 14080, A
C 41	42.4	1.2	23849	4 US-09-949-016-13813	Sequence 13813, A
C 42	42.4	1.2	39154	4 US-09-949-016-12384	Sequence 12384, A
C 43	42.4	1.2	39154	4 US-09-949-016-12801	Sequence 12801, A
C 44	42.4	1.2	39443	4 US-09-949-016-14326	Sequence 14326, A
C 45	42.4	1.2	39443	4 US-09-949-016-14327	Sequence 14327, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHEIFLINGER, F.
: APPLICANT: FALKNER, F. G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,463
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-9300
: TELEFAX: (703) 683-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pTZpc-F18
: US-08-232-463-14
Query Match 1.6%, Score 57.6, DB 1, Length 7218;

	Best Local Similarity	3.9%; Pred.No. 5.6e-05;
	Matches	15; Conservative 218; Mismatches 147; Indels 0; Gaps 0;
QY	510 TCTCTCCCATTCCAATTGACCGAATTAACAACGGCAGCAAAGCTCGACGTAGCCA	569
	: :	
Db	1078 YY	1137
QY	570 CAACACTCTAACACTATCTGCTCGCAGCAAAATGGCCGACCGTCTCTCTCGACTT	629
	: :	
Db	1138 YY	1197
QY	630 CAATTTCACCTCGATTCCTCTCTGCTGCTTGCCGCAACGCCACACGCACGACTCT	689
	: :	
Db	1198 YY	1257
QY	690 CGTCTCTGCGCACCACTATATGATTCGCCGAAGCACAATCCCATTTGATTT	749
	: :	
Db	1258 YY	1317
QY	750 CTACACAGTATTAGAGCTCAACACATTTCTTAACCGATGAATCAGAAGACATTGCA	809
	: :	
Db	1318 YY	1377
QY	810 AGCTAGGTTTCGAAACCGCCGCAATTCGGTTTACGCGACGACGCTTATACGCCGAG	869
	: :	
Db	1378 YYYGTA	1437
QY	870 ACAGATTCCTCAAGCTGCTT	889
Db	1438 CCAATTCTCTATCTCTTT	1457

```

RESULT 2
US-09-806-708B-22
; Sequence 22: Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: Consensus sequence of A.t., L.a., and B.n. PAB1 promoters
; US-09-806-708B-22

```

	Query Match	1.5#;	Score 54;	DB 4;	Length 1141;
	Best Local Similarity	10.0#;	Pred. No. 0.00016;		
	Matches	79;	Conservative 311;	Mismatches 389;	Indels 8; Gaps 2;
Qy	2393	TTAAGTAGAGATCCCTGGTGGTAACATGATGAGCGGTGATGGTGAACCTGGTGCTTTATTG	2452		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Db	67	KYRRWNNKSKRWMKWKKKKYBCANNTSBRVHYHARMXDMKTAYBMPTTNNGKGTGRHR	126		
Qy	2453	CAGAAGCTGTAAAGCCCTCTGAA-AACTTTGAAAATAATGATTATTCGAATTCAGACTGGG	2511		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Db	127	YWRRRAABDVTDDHHYVTVAMNNAWMTTCMDDDDKRTTRMMKKNNATGWDDTKRHAWN	186		
Qy	2512	GTCRCAGAGATGACGTTGATGAATACTGTGAAAAATGCCGTGCATATGTAAAG	2571		
	:	: : : :	:	:	:
	:	: : : :	:	:	:
Db	187	NGBETVVMWRKYIKTDROMSBKRMNYGNBWMGMSEVDVTVYWWVVDDMCCKRVRNRWATRG	246		
Qy	2572	GAGCAAGATGTGAAGATCCTAGCTGCTGGTGGCAATTGGCATGATTTCACTGTTCAQC	2631		

Dd 247 RMRUYVMAWBTFAHRRRNNYNGWTBAAVYCRRTWNNNNNNAKAKCKEAKYGYMRABVNSTC 306

Oy 2632 CAGAAGTATTTCTTAAAAAGCGAGTCATTCTTCAACGGAAGATATGGTTCTCTATG 2691

Dd 307 TTWMSKTTRVTSCVANNCRAGDARKDKHKMMWSAMGYNNNNNNNNNTYYTKAARBAHW 366

Oy 2692 GAATGTGATGCGCTACCATAGTGATGTGAATTAATCATGCATATTTCCATATCTGCATTG 2751

Dd 367 DMVWHSARKMHANAAHYSRKKMTBYEKXKTYVNNNNGTTTWHRMYAMWKMDMBGTVN 426

Oy 2752 CTCAAAATAATGCTGTCTTTGTGAGCTAAGAACAATAGTCCCCTAATACATGTCCCAA 2811

Dd 427 NNNNGGRYYMTKNKKMMYYKWRANNCMKWAMDHKCTCHNNTTYWWKKKYTNMNNCYKS 486

Oy 2812 AGTTGACCAGATTAACAAGTGTGCTGAGTAATATCTACTAATTAATGCT-----GCTT 2864

Dd 487 MTNGSKSHBBAAAVTYWMMWRRYAHANNNDYMWKACTWIKYBCSKMNNYAMAYRK 546

Oy 2865 GAAITTTTGAATCAAACTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTG 2924

Dd 547 SSMWTSYYWMKTNNSWRSRSDTSMGRANNYAABHHGYKMNRWBMWSHTWBHBAG 606

Oy 2925 AATAAGAGGAGTTGAGATGCGCTTAGTGCGTGTGCCAAGTTTTCTTCCCTGATT 2984

Dd 607 AAHTYMBMTBAKCHCMQWYKAKYACAGSGSNNNNNNNNNNNNNNNNNATTCARDVYAAS 666

Oy 2985 TTTTTCCTTTCGATTTAGGGTCAGTCAGAGCTGACGATTCAGAGACATTCACAGATGS 3044

Dd 667 RMYMANAKWYYXXEPAANNAAYTHANNWGCMNNATIDPRTRYKNNNNNNAGTGRKNNNN 726

Oy 3045 ATGTCAGAGCTGCAGAGATAATAGTATCAAGTGCAGAGATTAAGTCTCGGCTTTTG 3104

Dd 727 NAKNBSAAKYNALAAVKLAKGHMRYANKMARGMADAAABETTDKRNGCAYIKYTTNNN 786

Oy 3105 GGCGTCATACCGCATAGAAATGTTTACAGAGSTGAGGGAATTAATCTPACAAATTCATCA 3164

Dd 787 NTYRGVVNTPTAADMGANNNNNNNNNNNNNNNGMSDMWYTWMAYAYGTUNNNNNNNNNAY 846

Oy 3165 ATTGCT 3171

Dd 847 AMWTTWK 853

```

1      RESULT 3
2      US-09-949-016-11796/C
3      ; Sequence 11796, Application US/09949016
4      ; Patent No. 6812339
5      ; GENERAL INFORMATION:
6      ; APPLICANT: VENTER, J. Craig et al.
7      ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
8      ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
9      ; FILE REFERENCE: C001307
10     ; CURRENT APPLICATION NUMBER: US/09/949,016
11     ; CURRENT FILING DATE: 2000-04-14
12     ; PRIOR APPLICATION NUMBER: 60/241,755
13     ; PRIOR FILING DATE: 2000-10-20
14     ; PRIOR APPLICATION NUMBER: 60/237,768
15     ; PRIOR FILING DATE: 2000-10-03
16     ; PRIOR APPLICATION NUMBER: 60/231,498
17     ; PRIOR FILING DATE: 2000-09-08
18     ; NUMBER OF SEQ ID NOS: 207012
19     ; SOFTWARE: FastSeq for Windows Version 4.0
20     ; SEQ ID NO 11796
21     ; LENGTH: 57280
22     ; TYPE: DNA
23     ; ORGANISM: Human
24     US-09-949-016-11796

```

Query Match	1.4%	Score	50.2	DB	4	Length	57280
Best Local Similarity	53.9%	Pred. No.	0.031				
Matches	103	Conservative	0	Mismatches	88	Indels	0
						Gaps	0

QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCCAAGCCACGTCCTCCATCCCAT 744
DB 12981 ACAACACCACTCTACTATGATGATCTTCACCATCAATTAACATCACTTCACACCAAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACA 12911

RESULT 7
US-09-949-016-13542/c
Sequence 13542, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTNER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13542
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-13542

Query Match 1.4%; Score 50.2; DB 4; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACCTCTCAATCTATGTCGCCAAGAAATGGCGACCGTCTCTCTCC 624
DB 13101 AGCACATCAACCAACCACTGCTCCACCAACCACTTCACCATTAATGAGCTCC 13042
QY 625 GACTCAATTTCACTCCGATTCCTCTCTCTCTCTGCGACCGCCACACACCGCC 684
DB 13041 ACATCAATATCCCTATGATGATGTCGCCAAGCCACGCTTCACCAACCAAC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCCAAGCCACGTCCTCCATCCCAT 744
DB 12981 ACAACACCACTCTACTATGATGATCTTCACCATCAATTAACATCACTTCACACCAAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACA 12911

RESULT 8
US-09-949-016-13543/c
Sequence 13543, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTNER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13543
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-13543

Query Match 1.4%; Score 50.2; DB 4; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACCTCTCAATCTATGTCGCCAAGAAATGGCGACCGTCTCTCTCC 624
DB 13101 AGCACATCAACCAACCACTGCTCCACCAACCACTTCACCATTAATGAGCTCC 13042
QY 625 GACTCAATTTCACTCCGATTCCTCTCTCTCTCTGCGACCGCCACACACCGCC 684
DB 13041 ACATCAATATCCCTATGATGATGTCGCCAAGCCACGCTTCACCAACCAAC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCCAAGCCACGTCCTCCATCCCAT 744
DB 12981 ACAACACCACTCTACTATGATGATCTTCACCATCAATTAACATCACTTCACACCAAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACA 12911

RESULT 9
US-09-949-016-13544/c
Sequence 13544, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTNER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13544
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-13544

Query Match 1.4%; Score 50.2; DB 4; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACCTCTCAATCTATGTCGCCAAGAAATGGCGACCGTCTCTCTCC 624
DB 13101 AGCACATCAACCAACCACTGCTCCACCAACCACTTCACCATTAATGAGCTCC 13042
QY 625 GACTCAATTTCACTCCGATTCCTCTCTCTCTCTGCGACCGCCACACACCGCC 684
DB 13041 ACATCAATATCCCTATGATGATGTCGCCAAGCCACGCTTCACCAACCAAC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCCAAGCCACGTCCTCCATCCCAT 744
DB 12981 ACAACACCACTCTACTATGATGATCTTCACCATCAATTAACATCACTTCACACCAAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACA 12911

RESULT 10

US-09-949-016-13545/C
; Sequence 13545, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13545
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13545

Query Match 1.4%; Score 50.2; DB 4; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACCTTACATATCTGCTCCGCGCAAGATGGCGGACCGCTTCTCTCC 624
DB 13101 AGCACCATCAACACCAACAGCTGCTCCACCACTACCACTTACCATTAAGCTCC 13042
QY 625 GACTTCAATTTACCTCGATTCCTCTCTCTCTCTGCGCACGCGCACCAACGCGC 684
DB 13041 ACATCAATATCCCTATCAGCTCCACCAACCGCTTTCACATCAGCACCAACGCGC 12982
QY 685 ACTCTGCTCTCTGCGCACCATTAATGATGTGCGGAGCGGACGTCCTCCATTC 744
DB 12981 ACAACCACTCACTCATATGACTTCCACCATCAACATTAACATTCACCAACAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACGA 12911

RESULT 11

US-09-949-016-14633/C
; Sequence 14633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14633
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14633

Query Match 1.4%; Score 50.2; DB 4; Length 57280;

Best Local Similarity 53.9%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACCTTACATATCTGCTCCGCGCAAGATGGCGGACCGCTTCTCTCC 624
DB 13101 AGCACCATCAACACCAACAGCTGCTCCACCACTACCACTTACCATTAAGCTCC 13042
QY 625 GACTTCAATTTACCTCGATTCCTCTCTCTCTCTGCGCACGCGCACCAACGCGC 684
DB 13041 ACATCAATATCCCTATCAGCTCCACCAACCGCTTTCACATCAGCACCAACGCGC 12982
QY 685 ACTCTGCTCTCTGCGCACCATTAATGATGTGCGGAGCGGACGTCCTCCATTC 744
DB 12981 ACAACCACTCACTCATATGACTTCCACCATCAACATTAACATTCACCAACAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACGA 12911

RESULT 12

US-09-949-016-14634/C
; Sequence 14634, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14634
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14634

Query Match 1.4%; Score 50.2; DB 4; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACCTTACATATCTGCTCCGCGCAAGATGGCGGACCGCTTCTCTCC 624
DB 13101 AGCACCATCAACACCAACAGCTGCTCCACCACTACCACTTACCATTAAGCTCC 13042
QY 625 GACTTCAATTTACCTCGATTCCTCTCTCTCTCTGCGCACGCGCACCAACGCGC 684
DB 13041 ACATCAATATCCCTATCAGCTCCACCAACCGCTTTCACATCAGCACCAACGCGC 12982
QY 685 ACTCTGCTCTCTGCGCACCATTAATGATGTGCGGAGCGGACGTCCTCCATTC 744
DB 12981 ACAACCACTCACTCATATGACTTCCACCATCAACATTAACATTCACCAACAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACGA 12911

RESULT 13

US-09-949-016-14635/C
; Sequence 14635, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

1 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
2 FILE REFERENCE: CL001307
3 CURRENT APPLICATION NUMBER: US/09/949,016
4 CURRENT FILING DATE: 2000-04-14
5 PRIOR APPLICATION NUMBER: 60/241,755
6 PRIOR FILING DATE: 2000-10-20
7 PRIOR APPLICATION NUMBER: 60/237,768
8 PRIOR FILING DATE: 2000-10-03
9 PRIOR APPLICATION NUMBER: 60/231,498
10 PRIOR FILING DATE: 2000-09-08
11 NUMBER OF SEQ ID NOS: 207012
12 SOFTWARE: FastSeq for Windows Version 4.0
13 SEQ ID NO 14635
14 LENGTH: 57280
15 TYPE: DNA
16 ORGANISM: Human
17 US-09-949-016-14635

Query Match 1.4%; Score 50.2; DB 4; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACACCTCTACATATCTGCTCCGCCAGCAATGGGCGGACCGCTTCTCTCC 624
DB 13101 AGCACCATCAACACCACTGCTCTCCACCACTTACCATTTATGAGCTCC 13042
QY 625 GACTCAATTGACCTCCGATTCCTCTCTCTCTGCGCAGCCGACCAACCGCC 684
DB 13041 ACCATCAATTCCTTATGAGTCCACCAACCGCTTTCACCATCAACCAACCGC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGCTCCGAAAGCCAGTCCCATCCCAT 744
DB 12981 ACAACACCACTCATCTATGAGTCCACCATCAATTAACATCACTCCACCAAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACA 12911

RESULT 14
US-09-949-016-14636/c
1 Sequence 14636, Application US/09949016
2 Patent No. 6812339
3 GENERAL INFORMATION:
4 APPLICANT: VENTER, J. Craig et al.
5 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
6 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
7 FILE REFERENCE: CL001307
8 CURRENT APPLICATION NUMBER: US/09/949,016
9 CURRENT FILING DATE: 2000-04-14
10 PRIOR APPLICATION NUMBER: 60/241,755
11 PRIOR FILING DATE: 2000-10-20
12 PRIOR APPLICATION NUMBER: 60/237,768
13 PRIOR FILING DATE: 2000-10-03
14 PRIOR APPLICATION NUMBER: 60/231,498
15 PRIOR FILING DATE: 2000-09-08
16 NUMBER OF SEQ ID NOS: 207012
17 SOFTWARE: FastSeq for Windows Version 4.0
18 SEQ ID NO 14636
19 LENGTH: 57280
20 TYPE: DNA
21 ORGANISM: Human
22 US-09-949-016-14636

Query Match 1.4%; Score 50.2; DB 4; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACACCTCTACATATCTGCTCCGCCAGCAATGGGCGGACCGCTTCTCTCC 624
DB 13101 AGCACCATCAACACCACTGCTCTCCACCACTTACCATTTATGAGCTCC 13042
QY 625 GACTCAATTGACCTCCGATTCCTCTCTCTCTGCGCAGCCGACCAACCGCC 684

DB 13041 ACCATCAATTCCTTATGAGTCCACCAACCGCTTTCACCATCAACCAACCGC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGCTCCGAAAGCCAGTCCCATCCCAT 744
DB 12981 ACAACACCACTCATCTATGAGTCCACCACTCAATTAACATCACTCCACCAAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACA 12911

RESULT 15
US-09-949-016-14637/c
1 Sequence 14637, Application US/09949016
2 Patent No. 6812339
3 GENERAL INFORMATION:
4 APPLICANT: VENTER, J. Craig et al.
5 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
6 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
7 FILE REFERENCE: CL001307
8 CURRENT APPLICATION NUMBER: US/09/949,016
9 CURRENT FILING DATE: 2000-04-14
10 PRIOR APPLICATION NUMBER: 60/241,755
11 PRIOR FILING DATE: 2000-10-20
12 PRIOR APPLICATION NUMBER: 60/237,768
13 PRIOR FILING DATE: 2000-10-03
14 PRIOR APPLICATION NUMBER: 60/231,498
15 PRIOR FILING DATE: 2000-09-08
16 NUMBER OF SEQ ID NOS: 207012
17 SOFTWARE: FastSeq for Windows Version 4.0
18 SEQ ID NO 14637
19 LENGTH: 57280
20 TYPE: DNA
21 ORGANISM: Human
22 US-09-949-016-14637

Query Match 1.4%; Score 50.2; DB 4; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.031;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACACCTCTACATATCTGCTCCGCCAGCAATGGGCGGACCGCTTCTCTCC 624
DB 13101 AGCACCATCAACACCACTGCTCTCCACCACTTACCATTTATGAGCTCC 13042
QY 625 GACTCAATTGACCTCCGATTCCTCTCTCTCTGCGCAGCCGACCAACCGCC 684
DB 13041 ACCATCAATTCCTTATGAGTCCACCAACCGCTTTCACCATCAACCAACCGC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGCTCCGAAAGCCAGTCCCATCCCAT 744
DB 12981 ACAACACCACTCATCTATGAGTCCACCATCAATTAACATCACTCCACCAAC 12922
QY 745 GATTCTACCA 755
DB 12921 CATCATCACA 12911

Search completed: June 10, 2005, 01:02:51
Job time : 583 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 18:54:19 ; Search time 1797 Seconds
(without alignments)
12079.962 Million cell updates/sec

Title: US-10-600-070-3

Perfect score: 3667

Sequence: 1 tgcctcgcataagagagat.....ctataacataagggctacaa 3667

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3667	100.0	3667	12	ADJ38130 Arabidops
2	3663.8	99.9	3667	12	ADJ38136 Arabidops
3	1931	52.7	2637	12	ADJ38210 Placitid d
4	1724	47.0	2406	12	ADJ38129 Arabidops
5	1720.8	46.9	2406	12	ADJ38135 Arabidops
6	1717.6	46.8	2406	12	ADJ38208 Placitid d
7	402.2	11.0	561	12	ADJ38212 Placitid d
8	295.6	8.1	631	12	ADJ38264 Placitid d
9	208.8	5.7	660	12	ADJ38215 Placitid d
10	194	5.3	2283	12	ADJ38206 Placitid d
11	178.6	4.9	607	13	ACN48781 Cotton pr
12	175	4.8	439	12	ADP93143 Cotton ex
13	161	4.4	545	12	ADJ38255 Placitid d
14	159.4	4.3	552	13	ACM48855 Cotton pr
15	158	4.3	608	12	ADJ38217 Placitid d
16	145.6	4.0	527	12	ADJ38214 Placitid d
17	136.2	3.7	647	12	ADJ38265 Placitid d
18	134.6	3.7	652	12	ADJ38266 Placitid d
19	132.2	3.6	563	12	ADJ38258 Placitid d
20	130.2	3.6	446	12	ADJ38231 Placitid d

21	127.6	3.5	479	12	ADJ38230	Adj38230 Placitid d
22	125	3.4	537	12	ADJ38223	Adj38223 Placitid d
23	122.6	3.3	307	12	ADJ38218	Adj38218 Placitid d
24	118.6	3.2	871	12	ADJ38232	Adj38232 Placitid d
25	93.6	2.6	491	12	ADJ38254	Adj38254 Placitid d
26	92.6	2.5	187	12	ADJ38216	Adj38216 Placitid d
27	92	2.5	418	12	ADJ38224	Adj38224 Placitid d
28	91.4	2.5	460	12	ADJ38256	Adj38256 Placitid d
29	90.4	2.5	420	12	ADJ38227	Adj38227 Placitid d
30	89.4	2.4	604	12	ADJ38237	Adj38237 Placitid d
31	89.4	2.4	653	12	ADJ38228	Adj38228 Placitid d
32	87.2	2.4	480	12	ADJ38225	Adj38225 Placitid d
33	86.8	2.4	622	12	ADJ38226	Adj38226 Placitid d
34	85.2	2.3	300	12	ADJ38260	Adj38260 Placitid d
35	81.8	2.2	360	12	ADJ38259	Adj38259 Placitid d
36	80.8	2.2	606	12	ADJ38257	Adj38257 Placitid d
37	78	2.1	416	12	ADJ38219	Adj38219 Placitid d
38	77.2	2.1	549	12	ADJ38261	Adj38261 Placitid d
39	75.8	2.1	549	12	ADJ38233	Adj38233 Placitid d
40	66.6	1.8	336	12	ADJ38222	Adj38222 Placitid d
41	65.6	1.8	309	12	ADJ38221	Adj38221 Placitid d
42	60.2	1.6	295	12	ADJ38213	Adj38213 Placitid d
43	59.4	1.6	400	12	ADJ38263	Adj38263 Placitid d
44	54	1.5	535	12	ADJ38229	Adj38229 Placitid d
45	51.6	1.4	2000	8	ADA71938	Ada71938 Rice Gene

ALIGNMENTS

RESULT 1	ADJ38130	ADJ38130 standard; DNA; 3667 BP.
XX	ADJ38130;	
XX	AC	06-MAY-2004 (first entry)
XX	DT	
XX	DE	Arabidopsis thaliana Arc6-1 genomic DNA SegID2.
XX	KW	prokaryotic type; plastid division; Ftn2; ARCE; ARC5; Fzo; plant cell;
XX	KM	agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX	XX	herbicide target; ds.
OS	Arabidopsis thaliana.	
XX	XX	WO2004001003-A2.
PN	XX	31-DEC-2003.
PD	XX	
XX	PF	20-JUN-2003; 2003WO-US019536.
XX	PR	20-JUN-2002; 2002US-0390140P.
PR	09-AUG-2002; 2002US-040242P.	
PR	20-JUN-2003; 2003US-00600070.	
XX	PA	(UNMS) UNIV MICHIGAN STATE.
XX	PI	Osteryoung KM, Vitha S, Koksharova OA, Gao H;
DR	XX	WPI; 2004-082486/08.
DR	P-PSDB; ADJ38202.	
PT	XX	New isolated Ftn2, ARCE and/or Fzo-like nucleic acid sequences, useful
PT	XX	for further characterizing plastid division in plant cells, and in
PT	XX	varying agronomic and horticultural characteristics of economically
PT	XX	important plants.
XX	XX	Claim 1; SEQ ID NO 2; 287pp; English.
PS	XX	This invention relates to novel prokaryotic type or plastid division and
CC	XX	related genes and proteins. In particular, the invention relates to novel
CC	XX	Ftn2 (ARCE), ARC5 and Fzo-like genes and polypeptides. The methods and

CC compositions of the present invention are useful for further
CC characterizing plant/division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a DNA sequence
CC which is related to the invention.

XX Sequence 3667 BP; 983 A; 670 C; 842 G; 1172 T; 0 U; 0 Other;

Query Match 100.0%; Score 3667; DB 12; Length 3667;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TGTTCGATTAAGAGAAATACAAATTAAGCAATTTGCTTGAATTTCAACAAGATTTTG 60
DB 1 TGTTCGATTAAGAGAAATACAAATTAAGCAATTTGCTTGAATTTCAACAAGATTTTG 60
QY 61 CTGGCTAATAGAGATTCATTTGGCTCTGTTTCTTTTACATTAATGATTTTGG 120
DB 61 CTGGCTAATAGAGATTCATTTGGCTCTGTTTCTTTTACATTAATGATTTTGG 120
QY 121 AATTTACACATTCAGTTGATGTTAAGAAAGAGAGAAATGATGGGTTTGTGG 180
DB 121 AATTTACACATTCAGTTGATGTTAAGAAAGAGAGAAATGATGGGTTTGTGG 180
QY 181 TTTAACTTTAAAGTCAAGAAATTAAGTCAATGCTTATGCTATATGTAA 240
DB 181 TTTAACTTTAAAGTCAAGAAATTAAGTCAATGCTTATGCTATATGTAA 240
QY 241 AATGAAGCACTCCAAAGCTTTAGTGAATGATTAATTAAGATTAATCAT 300
DB 241 AATGAAGCACTCCAAAGCTTTAGTGAATGATTAATTAAGATTAATCAT 300
QY 301 AAAGTCCGTGGCGACTGAATCATATGATTTTATTTTTCAGTGTGAT 360
DB 301 AAAGTCCGTGGCGACTGAATCATATGATTTTATTTTTCAGTGTGAT 360
QY 361 GTTTTGTATTTAACTTATACCTCAAAATTCATTAACCTTACAGCAAAA 420
DB 361 GTTTTGTATTTAACTTATACCTCAAAATTCATTAACCTTACAGCAAAA 420
QY 421 CAGCTCTTCAATATGTAAGAAAGAAAGTTTGTAGAGCTTAAAGAACTCCC 480
DB 421 CAGCTCTTCAATATGTAAGAAAGAAAGTTTGTAGAGCTTAAAGAACTCCC 480
QY 481 ATGGAAGCTGAGTCAAGTGGATGATGCTCCCATTCGAATATGCGATTA 540
DB 481 ATGGAAGCTGAGTCAAGTGGATGATGCTCCCATTCGAATATGCGATTA 540
QY 541 CCGGCGAAGCAAAAGCTCCGAGCTAGGCAACAACCTCTACATCTGCTCCG 600
DB 541 CCGGCGAAGCAAAAGCTCCGAGCTAGGCAACAACCTCTACATCTGCTCCG 600
QY 601 AATGGGCGAAGCTCTTCTCTCGACTTCAATTTCACTCCGATTCCTCTCT 660
DB 601 AATGGGCGAAGCTCTTCTCTCGACTTCAATTTCACTCCGATTCCTCTCT 660
QY 661 TTGGCAACCGCAACAACCGCACTCTGCTCTGCGACCATGATGATGCC 720
DB 661 TTGGCAACCGCAACAACCGCACTCTGCTCTGCGACCATGATGATGCC 720
QY 721 GAAGCGACGTCCTCCATTCATTTTCAACAGGATTAAGAGCTCAAAACATTT 780
DB 721 GAAGCGACGTCCTCCATTCATTTTCAACAGGATTAAGAGCTCAAAACATTT 780
QY 781 TTAACCGATGATCAAGAGCAATGGAAGCTAGGTTTGAAGCCGCGCAATGG 840
DB 781 TTAACCGATGATCAAGAGCAATGGAAGCTAGGTTTGAAGCCGCGCAATGG 840
QY 841 TTGAGGACGACGCTTAAATGAGCGGAGACAGATTTCTCAAGCTGTGAAACT 900
DB 841 TTGAGGACGACGCTTAAATGAGCGGAGACAGATTTCTCAAGCTGTGAAACT 900
```

```
QY 901 TCTAATCTCGCTCTAGAGAGAGTACAAATGAAGGCTCTTGTATGATGAAGAGCTACA 960
DB 901 TCTAATCTCGCTCTAGAGAGAGTACAAATGAAGGCTCTTGTATGATGAAGAGCTACA 960
QY 961 GTATGACATGATGCTTGGAGTAAGGTAATTCGATTTCCGAAATATAAGTTCTTC 1020
DB 961 GTATGACATGATGCTTGGAGTAAGGTAATTCGATTTCCGAAATATAAGTTCTTC 1020
QY 1021 GTTTTAATTTCAATGATGATTAAGAGAAAGAACTTTATCTAGTGAAGGCTCTGGGC 1080
DB 1021 GTTTTAATTTCAATGATGATTAAGAGAAAGAACTTTATCTAGTGAAGGCTCTGGGC 1080
QY 1081 TCTCTGTATTCGAGAAAGGTGTGAGACTGAGATAGTCTTCGGGTGTGAGGCTCT 1140
DB 1081 TCTCTGTATTCGAGAAAGGTGTGAGACTGAGATAGTCTTCGGGTGTGAGGCTCT 1140
QY 1141 GCTTAAGAGAGGTGCTAAGTGTGTTAAGCAAGATGTGTTTGTATGAGGCTGTC 1200
DB 1141 GCTTAAGAGAGGTGCTAAGTGTGTTAAGCAAGATGTGTTTGTATGAGGCTGTC 1200
QY 1201 GTTTCTGATGCTCGAGGAGTCTATGCAATTCACCTGATTTTATCTGGTTA 1260
DB 1201 GTTTCTGATGCTCGAGGAGTCTATGCAATTCACCTGATTTTATCTGGTTA 1260
QY 1261 TGAATTTGTGAGGAAGCTTTGAAGCTTTTACAGTAAGTGTGCTTGTGTAATTTG 1320
DB 1261 TGAATTTGTGAGGAAGCTTTGAAGCTTTTACAGTAAGTGTGCTTGTGTAATTTG 1320
QY 1321 ACGAGGCTGCTTAAAGACCTTTCTGATTTGATCTTTGTATGATGCTTGTGTA 1380
DB 1321 ACGAGGCTGCTTAAAGACCTTTCTGATTTGATCTTTGTATGATGCTTGTGTA 1380
QY 1381 GGAGGAAGAGCAAGTAGCCTTGACCCGATTTACGTGCAAAATGATGAGACTTTGA 1440
DB 1381 GGAGGAAGAGCAAGTAGCCTTGACCCGATTTACGTGCAAAATGATGAGACTTTGA 1440
QY 1441 AGAGATCATCCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTGTGATGATTAAC 1500
DB 1441 AGAGATCATCCCGCTTATGCTTGAAGCTACTTGGCTTACCGCTGTGATGATTAAC 1500
QY 1501 TGGAAAGAGCTAAATGCTTAAAGCGGTGAGGGAATTTTGTGCTGTGAGAGAG 1560
DB 1501 TGGAAAGAGCTAAATGCTTAAAGCGGTGAGGGAATTTTGTGCTGTGAGAGAG 1560
QY 1561 TGGAGATCAGCTCTTGTGGGGGTTGACCCGTGAGAAATTAAGAGAGCGTTT 1620
DB 1561 TGGAGATCAGCTCTTGTGGGGGTTGACCCGTGAGAAATTAAGAGAGCGTTT 1620
QY 1621 ACGAATGACAGCTGTGAGCAAGTATACGTTTGAATCTTTTATTTTATTTTAC 1680
DB 1621 ACGAATGACAGCTGTGAGCAAGTATACGTTTGAATCTTTTATTTTATTTTAC 1680
QY 1681 TGAATTAATCTTAAGTTCATTTTAAATGATGTGTGAGAGTGAATCTTTTGA 1740
DB 1681 TGAATTAATCTTAAGTTCATTTTAAATGATGTGTGAGAGTGAATCTTTTGA 1740
QY 1741 GCTAACCCCAAGCAATATTCAGAGAGCTATTGAAGTTTACGAAGTTGCACTTGTCT 1800
DB 1741 GCTAACCCCAAGCAATATTCAGAGAGCTATTGAAGTTTACGAAGTTGCACTTGTCT 1800
QY 1801 GTGGCTCAAGCTTTTATGATTAAGAGAGCAACCTTTTACAGATCTGATTAAGCAATTC 1860
DB 1801 GTGGCTCAAGCTTTTATGATTAAGAGAGCAACCTTTTACAGATCTGATTAAGCAATTC 1860
QY 1861 CAGCAATTCAGAGCTAAGTAAGTAAGGCTATGGAATTCGAGAGTGTGATGATTA 1920
DB 1861 CAGCAATTCAGAGCTAAGTAAGTAAGGCTATGGAATTCGAGAGTGTGATGATTA 1920
QY 1921 CGGAATTAATTTGAGAGATTAAGCTCGCTTGAAGAGGAGCTGTGCACTGTTATAGGC 1980
DB 1921 CGGAATTAATTTGAGAGATTAAGCTCGCTTGAAGAGGAGCTGTGCACTGTTATAGGC 1980
QY 1981 AAAGTTGATGAATGCCGTAATGTGGTTGAGCTTGAAGAGATTCACAAATATAGGAAT 2040
```


DR WPI; 2004-082486/08.
DR P-PSDB; ADJ38203.
XX
PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Claim 1; SEQ ID NO 10; 287bp; English.
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a DNA sequence
CC which is related to the invention.
XX
SQ Sequence 3667 BP; 982 A; 669 C; 843 G; 1173 T; 0 U; 0 Other;
Query Match 99.9%; Score 3663.8; DB 12; Length 3667;
Best Local Similarity 99.9%; Freq. No. 0;
Matches 3665; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TGTTCTGCATTAAGAGAAATACAAATTAAGCAATTTGCTTGAATTTCAAGAAATTTG 60
DB 1 TGTTCTGCATTAAGAGAAATACAAATTAAGCAATTTGCTTGAATTTCAAGAAATTTG 60
QY 61 CTGGCTAATAGGATTCATTGCTCTGTTTCTTTTACATTTACATGTCATTAATAGTTTCG 120
DB 61 CTGGCTAATAGGATTCATTGCTCTGTTTCTTTTACATTTACATGTCATTAATAGTTTCG 120
QY 121 AATTTTCACTTACAGTGGATGTTAAGAAAGAGGGAAATTGAGGGGTTTGTGGG 180
DB 121 AATTTTCACTTACAGTGGATGTTAAGAAAGAGGGAAATTGAGGGGTTTGTGGG 180
QY 181 TTTAACTTTAAAGTAGTCAAGAAATTAAGTATTTGCTTGTCTATATGTGTA 240
DB 181 TTTAACTTTAAAGTAGTCAAGAAATTAAGTATTTGCTTGTCTATATGTGTA 240
QY 241 AATGAAGCAACTCCAAAGGTTCTTAAAGTGAATGATTATTAGACATTTAAATCAT 300
DB 241 AATGAAGCAACTCCAAAGGTTCTTAAAGTGAATGATTATTAGACATTTAAATCAT 300
QY 301 AAAATCCGTCGCTGTAATCATATGATTTTATTTTTCAGTGTGAT 360
DB 301 AAAATCCGTCGCTGTAATCATATGATTTTATTTTTCAGTGTGAT 360
QY 361 GTTTTGTGATTTAATTATCTACTCAAAATCAAAATTCATTAACCTGACACAAA 420
DB 361 GTTTTGTGATTTAATTATCTACTCAAAATCAAAATTCATTAACCTGACACAAA 420
QY 421 CAGTCTCTCAATATGTAAACAGAAAGTTTGTAGAGCTTAAAGACATCCC 480
DB 421 CAGTCTCTCAATATGTAAACAGAAAGTTTGTAGAGCTTAAAGACATCCC 480
QY 481 ATGAAGCTCTGATCACTGTGGCATTTGTCTCTCCCATTCCAATTAAGCCATTACCA 540
DB 481 ATGAAGCTCTGATCACTGTGGCATTTGTCTCTCCCATTCCAATTAAGCCATTACCA 540
QY 541 CCGGCGACGACAAAGCTCCGAGTAGCAACAACCTCTACAACTTATCTGTCGCGACG 600
DB 541 CCGGCGACGACAAAGCTCCGAGTAGCAACAACCTCTACAACTTATCTGTCGCGACG 600
QY 601 AAATGGGCGACGCTCTCTCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
DB 601 AAATGGGCGACGCTCTCTCTCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
QY 661 TTGGCCACCGGCAACAACCGGCACTCTGTCCTCTGACCAACATATATGATGCTGCC 720
DB 661 TTGGCCACCGGCAACAACCGGCACTCTGTCCTCTGACCAACATATATGATGCTGCC 720

QY 721 GAACGCCACGTCCTCCATCCCATTTGATTTCTACAGGATATTAGAGCTCAAAACATTTCC 780
DB 721 GAACGCCACGTCCTCCATCCCATTTGATTTCTACAGGATATTAGAGCTCAAAACATTTCC 780
QY 781 TTAACCGATGATCAGAAAGACCTTGAAGCTTAGGGTTTGAACCCCGGCAATTTGGT 840
DB 781 TTAACCGATGATCAGAAAGACCTTGAAGCTTAGGGTTTGAACCCCGGCAATTTGGT 840
QY 841 TTGACGACGACGCTTAAATCAGCCGAGACAGATTTCTTCAAGCTGTGCGAAACTCTG 900
DB 841 TTGACGACGACGCTTAAATCAGCCGAGACAGATTTCTTCAAGCTGTGCGAAACTCTG 900
QY 901 TCTAATCTCGCTCTAGAAAGATACATGAAGCTTCTTGAATGATGAAGAACTACA 960
DB 901 TCTAATCTCGCTCTAGAAAGATACATGAAGCTTCTTGAATGATGAAGAACTACA 960
QY 961 GTCATCACTGATGTTCTTGGGATAGGTAAATTTCCATTCGAAATATAAGTTTCTTC 1020
DB 961 GTCATCACTGATGTTCTTGGGATAGGTAAATTTCCGAAATATAAGTTTCTTC 1020
QY 1021 GTTTTAAATTTCAATGAATTTGATTAAGAGAGAACTTTATCTAGTGAAGTTCTGGGGC 1080
DB 1021 GTTTTAAATTTCAATGAATTTGATTAAGAGAGAACTTTATCTAGTGAAGTTCTGGGGC 1080
QY 1081 TCTCTGTGATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTCGGGTTGGTGAAGCTCT 1140
DB 1081 TCTCTGTGATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTCGGGTTGGTGAAGCTCT 1140
QY 1141 GCTTAAGAGAGGTGTGCTTAAGCTGTTTAAGCAAGATGTGGTTTAAATAGGCGCTTC 1200
DB 1141 GCTTAAGAGAGGTGTGCTTAAGCTGTTTAAGCAAGATGTGGTTTAAATAGGCGCTTC 1200
QY 1201 GTTTCGATGTCTCGAGGAGATGCTATGCAATGATCCACTGATTTTAACTGTGTA 1260
DB 1201 GTTTCGATGTCTCGAGGAGATGCTATGCAATGATCCACTGATTTTAACTGTGTA 1260
QY 1261 TGAATTTGTTGAGAGACCTTGAAGCTTTTACAGTAAAGTTGACTTGTGGTAATTTG 1320
DB 1261 TGAATTTGTTGAGAGACCTTGAAGCTTTTACAGTAAAGTTGACTTGTGGTAATTTG 1320
QY 1321 ACGAGCTGTGGCTTTAATAAATCTTTGATTTGAATCTTTGTATATGATGCTGTGTA 1380
DB 1321 ACGAGCTGTGGCTTTAATAAATCTTTGATTTGAATCTTTGTATATGATGCTGTGTA 1380
QY 1381 GAGAGAGAGAGCAAGTACCTTGACACCGAATTAACGTGCAAAATGATGAGACTTTGA 1440
DB 1381 GAGAGAGAGAGCAAGTACCTTGACACCGAATTAACGTGCAAAATGATGAGACTTTGA 1440
QY 1441 AGAGATCACTCCGCTTATGTCTTGAAGCTACTTGGCTTACCGCTTGTGATGATTAAGC 1500
DB 1441 AGAGATCACTCCGCTTATGTCTTGAAGCTACTTGGCTTACCGCTTGTGATGATTAAGC 1500
QY 1501 TCGGAAAAGCTAAATGTGTTAAAGCGGTGTGCGGAATATTTGTGTCTGTGGAAGAG 1560
DB 1501 TCGGAAAAGCTAAATGTGTTAAAGCGGTGTGCGGAATATTTGTGTCTGTGGAAGAG 1560
QY 1561 TGGAGCATCAGCTCTGTGGGGGTTGACCGGTGAAGATTATGAATAGAGGCTTTT 1620
DB 1561 TGGAGCATCAGCTCTGTGGGGGTTGACCGGTGAAGATTATGAATAGAGGCTTTT 1620
QY 1621 ACGAATGACAGCTGCTGAGAGAGGTATCAAGTTAGATCTTTTAAATTTCTTTAGCA 1680
DB 1621 ACGAATGACAGCTGCTGAGAGAGGTATCAAGTTAGATCTTTTAAATTTCTTTAGCA 1680
QY 1681 TGAATATACCTTAAGTTTCTCATTTTAAATGATGTGTGTGATGATCTTTTGTGA 1740
DB 1681 TGAATATACCTTAAGTTTCTCATTTTAAATGATGTGTGTGATGATCTTTTGTGA 1740
QY 1741 GCTAACCCAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTT 1800
DB 1741 GCTAACCCAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTT 1800

QY 1801 GTGGCTCAGGCTTTTATGTAAGAGCCACACTTTTACAGAGATGCTGATAGCAATTC 1860
 DB 1801 GTGGCTCAGGCTTTTATGTAAGAGCCACACTTTTACAGAGATGCTGATAGCAATTC 1860
 QY 1861 CAGCAACTTCAGCAGGCTTAAGGTAATGGCTATGAGATTCCTGGCATGTTGTATGATCA 1920
 DB 1861 CAGCAACTTCAGCAGGCTTAAGGTAATGGCTATGAGATTCCTGGCATGTTGTATGATCA 1920
 QY 1921 CGGATTAATTTGGGAGATGACTTGGTCTAAGAAAGGGGACTGTGCACTGCTTATAGGC 1980
 DB 1921 CGGATTAATTTGGGAGATGACTTGGTCTAAGAAAGGGGACTGTGCACTGCTTATAGGC 1980
 QY 1981 AAAGTTGATGATGCGCTATGATGTTGGGCTTGAACATGAGGATTCACATATAGGAT 2040
 DB 1981 AAAGTTGATGATGCGCTATGATGTTGGGCTTGAACATGAGGATTCACATATAGGAT 2040
 QY 2041 CCAAGTATTTGGAGTTGTTGTTGGAGAAATTCAAATCGTATGACATATGATATCTCCT 2100
 DB 2041 CCAAGTATTTGGAGTTGTTGTTGGAGAAATTCAAATCGTATGACATATGATATCTCCT 2100
 QY 2101 GGACATATGCAAAATTTGTAAGAAACCTGTTGGCAGGGGTTGCTTCTTACAGGTTGAGAC 2160
 DB 2101 GGACATATGCAAAATTTGTAAGAAACCTGTTGGCAGGGGTTGCTTCTTACAGGTTGAGAC 2160
 QY 2161 ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATCCATGCTTTTGAATTAC 2220
 DB 2161 ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATCCATGCTTTTGAATTAC 2220
 QY 2221 TTGGAAGAGTGAAGTATGTTCAAGGTTCTCTTTTAACTGCTGCTGCAACTATGCGAAG 2280
 DB 2221 TTGGAAGAGTGAAGTATGTTCAAGGTTCTCTTTTAACTGCTGCTGCAACTATGCGAAG 2280
 QY 2281 ATTGGAAGCCGAGCATGTAAGAGCTATGCTATGCGAGCATGCGAAGATTTTCTTCC 2340
 DB 2281 ATTGGAAGCCGAGCATGTAAGAGCTATGCTATGCGAGCATGCGAAGATTTTCTTCC 2340
 QY 2341 CGCTATACAGATAGAACTCGGCTGAACCCCAAGATGTCAGAGACAGTGTATTAGTGA 2400
 DB 2341 CGCTATACAGATAGAACTCGGCTGAACCCCAAGATGTCAGAGACAGTGTATTAGTGA 2400
 QY 2401 GATCCTGTTGTAACATATGAGGCGGTATGCTGATGCTGCTGCTTTTATGCGAAGCT 2460
 DB 2401 GATCCTGTTGTAACATATGAGGCGGTATGCTGATGCTGCTGCTTTTATGCGAAGCT 2460
 QY 2461 GTAAAGCCCTCTGAAACTTTTGAACTAATGATTATGCAATTCGAGCTGGGCTCTCAAG 2520
 DB 2461 GTAAAGCCCTCTGAAACTTTTGAACTAATGATTATGCAATTCGAGCTGGGCTCTCAAG 2520
 QY 2521 AGTAGCGTTGATGAACCTACTGTTGAATGTCCTGCTGATATGTTAAAGAGGCAAGT 2580
 DB 2521 AGTAGCGTTGATGAACCTACTGTTGAATGTCCTGCTGATATGTTAAAGAGGCAAGT 2580
 QY 2581 GTGAAGATCTAGCTGCTGCTGCTGCAATGCACTGATTCACCTGTCAGCCAGAGAT 2640
 DB 2581 GTGAAGATCTAGCTGCTGCTGCTGCAATGCACTGATTCACCTGTCAGCCAGAGAT 2640
 QY 2641 TTTCTTAAAGACAGCTCACTTTTCAAGCAAGATATGCTTCTTATGGAATCTGAT 2700
 DB 2641 TTTCTTAAAGACAGCTCACTTTTCAAGCAAGATATGCTTCTTATGGAATCTGAT 2700
 QY 2701 GTGCTACCATAGGATATTAATATGATGCAATTTTCAATATCTGCAATGCTCAAAATA 2760
 DB 2701 GTGCTACCATAGGATATTAATATGATGCAATTTTCAATATCTGCAATGCTCAAAATA 2760
 QY 2761 TGCTTGTGTTGAGCTAAGACATAGTTCACCTTAAATACATGTCCTCAAAAGTTGATCC 2820
 DB 2761 TGCTTGTGTTGAGCTAAGACATAGTTCACCTTAAATACATGTCCTCAAAAGTTGATCC 2820
 QY 2821 AAGATTAACAAGTGTCTGAGTAAATTTCACTAATATGCTGCTTGAATTTTGTATCAA 2880
 DB 2821 AAGATTAACAAGTGTCTGAGTAAATTTCACTAATATGCTGCTTGAATTTTGTATCAA 2880
 QY 2881 CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTGAATAACGATGATAG 2940

DB 2881 CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTAGAAATACATAGGATAG 2940
 QY 2941 AGATTGCTTATGCTGCTGCTTGTCCAACTTTTCTTCTGTAATTTTCTTTGATTT 3000
 DB 2941 AGATTGCTTATGCTGCTGCTTGTCCAACTTTTCTTCTGTAATTTTCTTTGATTT 3000
 QY 3001 AGGCTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTCAGA 3060
 DB 3001 AGGCTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGCTCAGA 3060
 QY 3061 GAATATGATATCCAAAGTGGCAAGAAATTAAGTCTCTGCTTTTGGGCTGATCAACCGAT 3120
 DB 3061 GAATATGATATCCAAAGTGGCAAGAAATTAAGTCTCTGCTTTTGGGCTGATCAACCGAT 3120
 QY 3121 AGAAATGTTACCAAGAGTGAAGGAAATTAATCTCAATTCATCAATGCTGGAATAACCTGT 3180
 DB 3121 AGAAATGTTACCAAGAGTGAAGGAAATTAATCTCAATTCATCAATGCTGGAATAACCTGT 3180
 QY 3181 TGGAATGATATATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
 DB 3181 TGGAATGATATATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
 QY 3241 GAATGCTGAAGATTTGGAATGCAAGAGCAAGCTGAACTGCCAGCTGGGTTGTTATG 3300
 DB 3241 GAATGCTGAAGATTTGGAATGCAAGAGCAAGCTGAACTGCCAGCTGGGTTGTTATG 3300
 QY 3301 ATTATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCAGCAATGGAACCCGTG 3360
 DB 3301 ATTATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCAGCAATGGAACCCGTG 3360
 QY 3361 CTGCTGGAAGCAACTCTGGAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420
 DB 3361 CTGCTGGAAGCAACTCTGGAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420
 QY 3421 ACAATGCTACTGATGTGAGAACTTACACAAAGATTAAGAAATTTCTGCTCAAGTCA 3480
 DB 3421 ACAATGCTACTGATGTGAGAACTTACACAAAGATTAAGAAATTTCTGCTCAAGTCA 3480
 QY 3481 GGTGGAATTAACCTGGAAGGCTCTGTTCTGTCATCATATATATCATATATGATGCTCT 3540
 DB 3481 GGTGGAATTAACCTGGAAGGCTCTGTTCTGTCATCATATATATATCATATATGATGCTCT 3540
 QY 3541 GAGCTTGGAGATTTCTTTGTTCTGTAATTTCTCTCTAAGTATGATGTTATTAATGA 3600
 DB 3541 GAGCTTGGAGATTTCTTTGTTCTGTAATTTCTCTCTAAGTATGATGTTATTAATGA 3600
 QY 3601 ACACAAAATAATTAACGTTCTTGGCACACCCCTTTCCTGATCTAATACATTAAGG 3660
 DB 3601 ACACAAAATAATTAACGTTCTTGGCACACCCCTTTCCTGATCTAATACATTAAGG 3660
 QY 3661 GCTACAA 3667
 DB 3661 GCTACAA 3667

RESULT 3
 ADJ38210 standard; DNA; 2637 BP.
 ID ADJ38210 standard; DNA; 2637 BP.
 XX ADJ38210;
 AC ADJ38210;
 AC ADJ38210;
 DT 06-MAY-2004 (first entry)
 XX 06-MAY-2004 (first entry)
 DE Plastid division-related Arc6 orthologue gene 3.
 XX Plastid division-related Arc6 orthologue gene 3.
 KW prokaryotic type; plastid division; Fun2; ARC6; ARC5; Pro; plant cell;
 KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KW herbicide target; gene; de.
 XX Arabidopsis thaliana.
 XX Arabidopsis thaliana.
 PN W02004001003-A2.

XX New Isolated Ftn2, ARCS and/or Pzo-like nucleic acid sequences, useful
PT for further characterizing plasmid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.

Claim 1; SEQ ID NO 1; 287bp; English.

XX This invention relates to novel prokaryotic type or plasmid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARCS), ARCS and Pzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plasmid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.

XX Sequence 2406 BP; 612 A; 492 C; 618 G; 684 T; 0 U; 0 Other;

Query Match 47.0%; Score 1724; DB 12; Length 2406;

Best Local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;

QY 481 ATGAGAGCTTGAATCAAGTCCGATGATGCTCTCCCATTCATATGCGCATACCA 540
DB 1 ATGAGAGCTTGAATCAAGTCCGATGATGCTCTCCCATTCATATGCGCATACCA 60
QY 541 CCGGCGAGCAAGAGCTCCGAGCGATGAGCAACACTCTACACTATGCTCGGACG 600
DB 61 CCGGCGAGCAAGAGCTCCGAGCGATGAGCAACACTCTACACTATGCTCGGACG 120
QY 601 AAATGGGCGGACGCTCTCTCTCCGACTTCAATTTCACTCCGATCTCTCTCTCTCC 660
DB 121 AAATGGGCGGACGCTCTCTCTCCGACTTCAATTTCACTCCGATCTCTCTCTCTCC 180
QY 661 TTGGCCACCGGACCAACCGGCACTCTGCTCTCTCTCCGACCATATGATGCTCC 720
DB 181 TTGGCCACCGGACCAACCGGCACTCTGCTCTCTCTCCGACCATATGATGCTCC 240
QY 721 GAAGCGACGCTCCGATGATGATTTCAAGAGATTAAGAGCTCAACCAATTTTC 780
DB 241 GAAGCGACGCTCCGATGATGATTTCAAGAGATTAAGAGCTCAACCAATTTTC 300
QY 781 TTAACCGATGAGATCAGAGAGCAATTCAGAGCTAGGTTTGAACCGCGCAATTCG 840
DB 301 TTAACCGATGAGATCAGAGAGCAATTCAGAGCTAGGTTTGAACCGCGCAATTCG 360
QY 841 TTCAAGCGACGAGCTTTAATCAAGCGGAGACAGATTTCTCAAGCTGCTTGAAATCTG 900
DB 361 TTCAAGCGACGAGCTTTAATCAAGCGGAGACAGATTTCTCAAGCTGCTTGAAATCTG 420
QY 901 TCTAATCTCGGCTTAGAGAGATTAAGAGATGATGATGATGATGATGATGATGATGAT 960
DB 421 TCTAATCTCGGCTTAGAGAGATTAAGAGATGATGATGATGATGATGATGATGATGAT 480
QY 961 GTCAATCACTGATGCTCTGAGAGATTAAGATTTGATTTGAGATTAAGATTTTTC 1020
DB 481 GTCAATCACTGATGCTCTGAGAGATTTTCTGAGATTTTCTGAGATTTTCTGAGAT 504
QY 1021 GTTTTAATTTCAAGATTAAGAGAGCACTTTAATCTAGTGAAGTTCTCGGAGC 1080
DB 505 -----AAGGTTCTCGGAGC 518
QY 1081 TCTCTGATTTGCAAGAGTGTGAGACTGATGATGTTCTTGAGGTTGTGAGGCTCT 1140
DB 519 TCTCTGATTTGCAAGAGTGTGAGACTGATGATGTTCTTGAGGTTGTGAGGCTCT 578
QY 1141 GCTTAAGAGAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 579 GCTTAAGAGAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 638
QY 1201 GTTCTGATGCTCGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1260

DB 639 GTTCTGATGCTCGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 698
QY 1261 TGAATTTGAGAGAGCTTTGAAGCTTTTACAGATTAAGATTTGATTTGATTTGATTTGAT 1320
DB 699 TGAATTTGAGAGAGCTTTGAGAGCTTTTAC----- 730
QY 1321 ACAGAGCTTGCTTTAAGAACTTTCTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1380
DB 731 -----A 731
QY 1381 GAGAGAGAGAGCAAGTACCTTGACCGGATTTACGTCACAAATGATGAGCTTTGGA 1440
DB 732 GAGAGAGAGAGCAAGTACCTTGACCGGATTTACGTCACAAATGATGAGCTTTGGA 791
QY 1441 AGAGATCACTCCGCTTATGCTTTGAGACTTATGAGCTTACCGCTTGATGATTTACG 1500
DB 792 AGAGATCACTCCGCTTATGCTTTGAGACTTATGAGCTTACCGCTTGATGATTTACG 851
QY 1501 TGGCAAAAGACTAATGTTTAAGCGGCTGCGAATATTTTGTGCTCTGTTGAGAGAG 1560
DB 852 TGGCAAAAGACTAATGTTTAAGCGGCTGCGAATATTTTGTGCTCTGTTGAGAGAG 911
QY 1561 TGGAGCACTCAGCTCTTGTGTTGGGTTTGAACCGGAGAAATTTATGAAATGAGGCTTTT 1620
DB 912 TGGAGCACTCAGCTCTTGTGTTGGGTTTGAACCGGAGAAATTTATGAAATGAGGCTTTT 971
QY 1621 ACGAATGACAGCTGCTGAGAGAGTATACAGTTATAGATACCTTTTAAATTTCTTTTGA 1680
DB 972 ACGAATGACAGCTGCTGAGAG----- 991
QY 1681 TGATTAATCTTAAGTTTCTCATTTTAATGATGTTGTGTTGATGTTGATCTTTTGTGA 1740
DB 992 -----AGTTGATCTTTTGTGA 1008
QY 1741 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAATTTGAGAGTGTGATCTTCT 1800
DB 1009 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAATTTGAGAGTGTGATCTTCT 1068
QY 1801 GTGGCTCAAGCTTTTATGTTAAGAGCCACACTTTTACAGATGCTGATTAAGCAATTC 1860
DB 1069 GTGGCTCAAGCTTTTATGTTAAGAGCCACACTTTTACAGATGCTGATTAAGCAATTC 1128
QY 1861 CAGCACTTCAAGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
DB 1129 CAGCACTTCAAGAGCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
QY 1921 CGGAATTAATGAGAGATGAGCTTCCGCTGAGAAAGGAGACTCTGTGCACTGTTATAGGC 1980
DB 1189 CGGAATTAATGAGAGATGAGCTTCCGCTGAGAAAGGAGACTCTGTGCACTGTTATAGGC 1248
QY 1981 AAAGTTGATGATGCTGATGTTGGGCTTGAACAGTGAAGATTCACAAATATAGGAAT 2040
DB 1249 AAAGTTGATGATGCTGATGTTGGGCTTGAACAGTGAAGATTCACAAATATAGGAAT 1308
QY 2041 CCAAGCTATTTGAGATTTGTTTGAAGATTTCAATTCGATGACATGATGATCTCCCT 2100
DB 1309 CCAAGCTATTTGAGATTTGTTTGAAGATTTCAATTCGATGACATGATGATCTCCCT 1368
QY 2101 GGAATATGAGAAATTTGTTGAAGCTGTTGAGAGGCTTCTTCTAGGTTCAAGAGC 2160
DB 1369 GGAATATGAGAAATTTGTTGAAGCTGTTGAGAGGCTTCTTCTAGGTTCAAGAGC 1428
QY 2161 ACCAAGATTAATAATTTAACTCGGAGACTAATGATGATGATGATGATGATGATGATGAT 2220
DB 1429 ACCAAGATTAATAATTTAACTCGGAGACTAATGATGATGATGATGATGATGATGATGAT 1488
QY 2221 TTGAAAAGAGTGAAGTGTGAGGCTTCTTGAAGCTCTGCTGCTGCACTATGCAAG 2280
DB 1489 TTGAAAAGAGTGAAGTGTGAGGCTTCTTGAAGCTCTGCTGCTGCACTATGCAAG 1548
QY 2281 ATTGAGCCGAGCATGTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340


```

Db      1549 ATTGAGCCGACGATGTAAGTAGTGTATGACAGGACCTGCAGAAAATTTCCTCC 1608
Qy      2341 CGGATATCAGATAGAACTGGCTGAACCCAGAGATGCGAAGACAGTGTAGTGA 2400
Db      1609 CGGATACAGATGAACTGGCTGAACCCAGAGATGCGAAGACAGTGTAGTGA 1668
Qy      2401 GATCTGTGTGTAACAAATGAGGCGTGAATGAGGCTGTGTCTTATATGACAGAACT 2460
Db      1669 GATCTGTGTGTAACAAATGAGGCGTGAATGAGGCTGTGTCTTATATGACAGAACT 1728
Qy      2461 GTAGAGCCCTGTGAAAATTGAAAATGATATGCAATTCAGAGCTGGGCTCTCAGAG 2520
Db      1729 GTAGAGCCCTGTGAAAATTGAAAATGATATGCAATTCAGAGCTGGGCTCTCAGAG 1788
Qy      2521 AGTAGCGTGTGTAACAACTAGTGAATGTCGCTGTGATGATGTAAGAGGCAAGT 2580
Db      1789 AGTAGCGTGTGTAACAACTAGTGAATGTCGCTGTGATGATGTAAGAGGCAAGT 1848
Qy      2581 GTGAAAGATCCTAGCTGTGTGTGCGCAATGGAATGCACTGATTCAGCCAGAGATAT 2640
Db      1849 GTGAAAGATCCTAGCTGTGTGTGCGCAATGGAATGCACTGATTCAGCCAGAGATAT 1908
Qy      2641 TTTCTTAAAGACAGCTCATCTTTCAACGCAAGATATGTTCTTCTATGAAATCTGAT 2700
Db      1909 TTTCTTAAAGACAGCTCATCTTTCAACGCAAGATATGTTCTTCTATGAAATCTGAT 1968
Qy      2701 GTGCTACCCATAGGTATGATTAATGATGCAATTTGATATGCAATGCTCAAAATA 2760
Db      1969 GTGCTACCA----- 1978
Qy      2761 TGCTGTGTTGTGAGTAAAGAACATAGTCCACTAATACATGTCACAAAATTGTACC 2820
Db      1979 ----- 1978
Qy      2821 AAGATTAACAAGTTGCTGAGTAAATTTCACTAATATGCTGTGAATTTTGTGACAA 2880
Db      1979 ----- 1978
Qy      2881 CTGTAGACAGAAATGTAAATTTCACTGCAACATTTCTGTTAGAAATAGGATTAG 2940
Db      1979 ----- 1978
Qy      2941 AGATGCGCTTAGTGTGCTTTGTCCAACTTTCTTCTGATTTTCTTTTCGATTT 3000
Db      1979 ----- 1978
Qy      3001 AGGCTCAGTCAAGCTGATGCAATTCAGAAACATTTCCCAAGATGATGCTAGACTGAGA 3060
Db      1980 AGGCTCAGTCAAGCTGATGCAATTCAGAAACATTTCCCAAGATGATGCTAGACTGAGA 2039
Qy      3061 GAATATGATATCCAAAGTGGCGAAGATTAATGCTGTGGCTTTGGGCTGATCACCCTAT 3120
Db      2040 GAATATGATATCCAAAGTGGCGAAGATTAATGCTGTGGCTTTGGGCTGATCACCCTAT 2099
Qy      3121 AGAAATGTTACAGAGGAGGAGGAATAATCTACATCAATCAATGTTGTGAAGACTGT 3180
Db      2100 AGAAATGTTACAG----- 2113
Qy      3181 TGACATGATTAATAGTGTGTGCTGTGTTGATTTCTGTTAATTAAGTGTGAGGGC 3240
Db      2114 -----AGGTTTGTGATGGGC 2128
Qy      3241 GAATGCTGAAGATTTGACATGACAGACAGCTGAACCTGCGCACTTGGGTGTTATAG 3300
Db      2129 GAATGCTGAAGATTTGACATGACAGACAGCTGAACCTGCGCACTTGGGTGTTATAG 2188
Qy      3301 ATTATACACTGTTGAACCTATCTGTGACAGTGTGACAGCTCTCAGCAGATGAAACCGGTG 3360
Db      2189 ATTATACACTGTTGAACCTATCTGTGACAGTGTGACAGCTCTCAGCAGATGAAACCGGTG 2248
Qy      3361 CTGTGTGGAAGCAACTCTGGAAGAGTCTGTGTTCTATCTGATTTGGTTGATCAGAAA 3420
Db      2249 CTGTGTGGAAGCAACTCTGGAAGAGTCTGTGTTCTATCTGATTTGGTTGATCAGAAA 2308

```

```

Qy      3421 ACAATGCTACTGATGTGAGAACTTACACAACAAGATGAAGTTTCTGTGCTCAAGTCAG 3480
Db      2309 ACAATGCTACTGATGTGAGAACTTACACAACAAGATGAAGTTTCTGTGCTCAAGTCAG 2368
Qy      3481 GGTGAAAAATCACTGAAGGCTCTGTCTTTCGATCATATA 3518
Db      2369 GGTGAAAAATCACTGAAGGCTCTGTCTTTCGATCATATA 2406

RESULT 5
ADJ38135
ID ADJ38135 standard; cDNA; 2406 BP.
XX
XX
AC ADJ38135;
XX
DT 06-MAY-2004 (first entry)
XX
DE Arabidopsis thaliana AtFtn2 cDNA SegID9.
XX
KW prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PP 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-040242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS ) UNIV MICHIGAN STATE.
XX
PI Oosteryoung KW, Vicha S, Koksharova OA, Gao H;
XX
DR WI: 2004-082486/08.
DR P-PSDB; ADJ38203.
XX
PT New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Claim 1; SEQ ID NO 9; 287bp; English.
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
SQ Sequence 2406 BP; 611 A; 491 C; 619 G; 685 T; 0 U; 0 Other;

Query Match 46.9%; Score 1720.8; DB 12; Length 2406;
Beet Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

Qy      481 ATGGAAGCTGTGAGTCAAGTGTGATGCTCTCCCAATTCGAATATGCGGATTAACA 540
Db      1 ATGGAAGCTGTGAGTCAAGTGTGATGCTCTCCCAATTCGAATATGCGGATTAACA 60

Qy      541 CCGGCGACGACAAAGCTCCGACGTAGCCACACACACTCTACAACTATCTGCTCGGCAGC 600
Db      61 CCGGCGACGACAAAGCTCCGACGTAGCCACACACACTCTACAACTATCTGCTCGGCAGC 120

```

QY 601 AAATGGGCGGACGCTTCTCTCCGACTTCAATTTCACCTCCGATTCTCTCTCTCC 660
DB 121 AAATGGGCGGACGCTTCTCTCCGACTTCAATTTCACCTCCGATTCTCTCTCTCC 180
QY 661 TTGGCCACCGGACCAACCAACGCGCACTCTCTCTCTCCGCACTTATGATGCTCC 720
DB 181 TTGGCCACCGGACCAACCAACGCGCACTCTCTCTCTCCGCACTTATGATGCTCC 240
QY 721 GAAGCGACGCTCCGATCCCATTTGATTCTTACAGAGGATTAGAGGCTCAAAACATTTT 780
DB 241 GAAGCGACGCTCCGATCCCATTTGATTCTTACAGAGGATTAGAGGCTCAAAACATTTT 300
QY 781 TTAAACGATGGAATCAGAAAGAGCATTCGAACTAGGGTTTCGAAACGCGCAATTCG 840
DB 301 TTAAACGATGGAATCAGAAAGAGCATTCGAACTAGGGTTTCGAAACGCGCAATTCG 360
QY 841 TTCAGCGACGACGCTTTAATCAAGCCGAGACAGATTCTTCAAGCTGCTTCCGAAATCTG 900
DB 361 TTCAGCGACGACGCTTTAATCAAGCCGAGACAGATTCTTCAAGCTGCTTCCGAAATCTG 420
QY 901 TCTAATCTCGGCTTAAAGAGAGTACAAATGAGGCTCTTGTATGATGAAGAGCTTACA 960
DB 421 TCTAATCTCGGCTTAAAGAGAGTACAAATGAGGCTCTTGTATGATGAAGAGCTTACA 480
QY 961 GTCATCACTGATGTTCTTGGGATAGGTAATTTGATTCGGAATATAAGTTCTTC 1020
DB 481 GTCATCACTGATGTTCTTGGGAT----- 504
QY 1021 GTTTTAATTTCAATATGATATAAGAGAGACTTTTATCTAGTGAAGTTCTTGGGAC 1080
DB 505 -----AAGTTCTTGGGAC 518
QY 1081 TCTCTGATTTTGCAGAAAGGTGAGTGAAGTATGTTCTTGGGTTGGAGGCTCT 1140
DB 519 TCTCTGATTTTGCAGAAAGGTGAGTGAAGTATGTTCTTGGGTTGGAGGCTCT 578
QY 1141 GCTTAAAGAGAGGTTGCTTAAGTCTTTAAAGCAAGATGTTGTTTAACTTATGCGCTTGC 1200
DB 579 GCTTAAAGAGAGGTTGCTTAAGTCTTTAAAGCAAGATGTTGTTTAACTTATGCGCTTGC 638
QY 1201 GTTCTCTGATGTTCTGAGGAGTGTATGAGTTCGATTCGATTTTATTAATCTGTTA 1260
DB 639 GTTCTCTGATGTTCTGAGGAGTGTATGAGTTCGATTCGATTTTATTAATCTGTTA 698
QY 1261 TGAAGTTTGGAGAGGCTTGAAGCTTTTACAGGTAAGTTTGAAGTTTGAATTTG 1320
DB 699 TGAAGTTTGGAGAGGCTTGAAGCTTTTAC----- 730
QY 1321 ACGAGCGTTGGCTTTATTAAGAACTTTCTGATTTGATACTTTGTATTTGAGTCTTGTGA 1380
DB 731 -----A 731
QY 1381 GAGAGAGAGAGCAAGTACCTTTCGACCGGATTTACGTGCACAAATTTGATGAGCTTTGGA 1440
DB 732 GAGAGAGAGAGCAAGTACCTTTCGACCGGATTTACGTGCACAAATTTGATGAGCTTTGGA 791
QY 1441 AGAGATCACTCCGCTTATGTTCTTGAAGTCTTGGCTTACCGCTTGGTATGATTTAGC 1500
DB 792 AGAGATCACTCCGCTTATGTTCTTGAAGTCTTGGCTTACCGCTTGGTATGATTTAGC 851
QY 1501 TGCAGAAAGACCTAAATGTTTAAAGCGGTGTCGGAATATTTTGTCTGTTGGAGAGG 1560
DB 852 TGCAGAAAGACCTAAATGTTTAAAGCGGTGTCGGAATATTTTGTCTGTTGGAGAGG 911
QY 1561 TGCAGCATCAAGCTTTTGTGGGGGTTTGAACCGTGAAGAGTTTATGATGAGCGGTTTT 1620
DB 912 TGCAGCATCAAGCTTTTGTGGGGGTTTGAACCGTGAAGAGTTTATGATGAGCGGTTTT 971
QY 1621 ACGAATACAGCTGTAGAGAGGATACAGTTTATGATACCTTTTAAATTTCTTTAGCA 1680
DB 972 ACGAATACAGCTGTAGAGC----- 991

QY 1681 TGAATATACTTAAAGTTTCTCATTTTAAATGATGTTGTGTTAGGTTGATCTTTTGTGA 1740
DB 992 -----AGTTTGAATCTTTTGTGA 1008
QY 1741 GCTAACCCAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACCTTGCCTT 1800
DB 1009 GCTAACCCAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACCTTGCCTT 1068
QY 1801 GTGGCTCAAGCTTTTATTTGTTAAGAAAGCCACACTTTTACAGAGTGTGATTAAGCAATTC 1860
DB 1069 GTGGCTCAAGCTTTTATTTGTTAAGAAAGCCACACTTTTACAGAGTGTGATTAAGCAATTC 1128
QY 1861 CAGCACTTCAAGAGCTTAAGGTAATGCTATGAGATTCCTGAGATTTGATGATACA 1920
DB 1129 CAGCACTTCAAGAGCTTAAGGTAATGCTATGAGATTCCTGAGATTTGATGATACA 1188
QY 1921 CGGAATTAATGGAGATGAGACTTCGCTGTGAAGAGGGAATCTGTGTGACGCTTATAGGC 1980
DB 1189 CGGAATTAATGGAGATGAGACTTCGCTGTGAAGAGGGAATCTGTGTGACGCTTATAGGC 1248
QY 1981 AAAGTTGATGAATGCGTATGTTGGGCTTGAACAGTGAAGATTCACAAATATAGGAAT 2040
DB 1249 AAAGTTGATGAATGCGTATGTTGGGCTTGAACAGTGAAGATTCACAAATATAGGAAT 1308
QY 2041 CCAAGCTATTTGAGATTTGTTTGAAGATTCAAATCTGATGACATGATGATCTCCCT 2100
DB 1309 CCAAGCTATTTGAGATTTGTTTGAAGATTCAAATCTGATGACATGATGATCTCCCT 1368
QY 2101 GGAATATGCAAAATTTGTTGGAACCTGGTTGGCAGGGGTTGCTTTCTAGGTTCAAGAC 2160
DB 1369 GGAATATGCAAAATTTGTTGGAACCTGGTTGGCAGGGGTTGCTTTCTAGGTTCAAGAC 1428
QY 2161 ACCAAAGATTAATAATTTAAATCTCGGGGACTACTATGATGATCTATGTTTGAATTAC 2220
DB 1429 ACCAAAGATTAATAATTTAAATCTCGGGGACTACTATGATGATCTATGTTTGAATTAC 1488
QY 2221 TTGAAAAGAGTGAAGGTAAGTTCAGGTTCTCTTTAGCTCTGCTGCACTATGAGCAAG 2280
DB 1489 TTGAAAAGAGTGAAGGTAAGTTCAGGTTCTCTTTAGCTCTGCTGCACTATGAGCAAG 1548
QY 2281 ATTGAGCGGAGATGGAAGGTAAGTTCAGGCACTGAGCAAGAAAGTTTCTCTCC 2340
DB 1549 ATTGAGCGGAGATGGAAGGTAAGTTCAGGCACTGAGCAAGAAAGTTTCTCTCC 1608
QY 2341 CGCTATACAGATGAATACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTATAGTGA 2400
DB 1609 CGCTATACAGATGAATACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTATAGTGA 1668
QY 2401 GATCTGTTGTTGAACAATGTAAGCCGTTGATGAGCTGTGTCTTTATTTGCAAGAGCT 2460
DB 1669 GATCTGTTGTTGAACAATGTAAGCCGTTGATGAGCTGTGTCTTTATTTGCAAGAGCT 1728
QY 2461 GTAAGACCTCTGAAAACTTTGAAACTTAATGATTAAGCAATTTGAGCTGGGGTCTCGAG 2520
DB 1729 GTAAGACCTCTGAAAACTTTGAAACTTAATGATTAAGCAATTTGAGCTGGGGTCTCGAG 1788
QY 2521 AGTAGCGTTGATGAATACTACTGTTGAAGTCCGTTGCTGATATGTTAAAGAGGCAAGT 2580
DB 1789 AGTAGCGTTGATGAATACTACTGTTGAAGTCCGTTGCTGATATGTTAAAGAGGCAAGT 1848
QY 2581 GTGAAGATCTTACGCTGTGTTGAGCAATTTGACTGATTTCACTGTTCAAGCAGAGAT 2640
DB 1849 GTGAAGATCTTACGCTGTGTTGAGCAATTTGACTGATTTCACTGTTCAAGCAGAGAT 1908
QY 2641 TTTCTTAAAGCAGCTCATCTTTTCAAGCAGAGATATGTTCTTATGGAATCGAT 2700
DB 1909 TTTCTTAAAGCAGCTCATCTTTTCAAGCAGAGATATGTTCTTATGGAATCGAT 1968
QY 2701 GTGCTACCAATAGGTAATGATTAATGATGCAATTTTCAATATCTGATTCACAAATA 2760
DB 1969 GTGCTACCA----- 1978
QY 2761 TGCTTGTTTTGTAGCTTAAGACATAGTTCCACTTAATATCATGTCCCAAAAGTTGTACC 2820

```

Db 1979 ----- 1978
QY 2821 AAGATTACAGATTGCTGAGTAATAATTCTAATTATGCTGTAATTTTTCATCAAA 2880
Db 1979 ----- 1978
QY 2881 CTGTAGACAGAAATGTAATTTCTACTCTCAACATTTCTGTTAGATAAGTAGATTAG 2940
Db 1979 ----- 1978
QY 2941 AGATTGCTTAGTGTGCTTGTTCACATTTCTTCTTGATTTTCTTTTCATTT 3000
Db 1979 ----- 1979
QY 3001 AGGCTCATGACAGCTGACATTCAGAAAGCATTTCCAGAAATGATGCTAGACTGACA 3060
Db 1980 AGGCTCATGACAGCTGACATTCAGAAAGCATTTCCAGAAATGATGCTAGACTGACA 2039
QY 3061 GAATATGATTCACAGTGGGAGAGATTAGTCTGTGCTTTTGGGCTGATCCCGCAT 3120
Db 2040 GAATATGATTCACAGTGGGAGAGATTAGTCTGTGCTTTTGGGCTGATCCCGCAT 2099
QY 3121 AGAATGTTTACAGAGGTGAGGAAATTAATCTACAATTCATCAATTGTGTGAAAACCTGT 3180
Db 2100 AGAATGTTTACAG----- 2113
QY 3181 TGGACATGATTAATAGTCTGTGCTTGTGATTTGATTATTAATAGTTTGGATGGGC 3240
Db 2114 -----AGGTTTGGATGGGC 2128
QY 3241 GAATGCTGAAGATTTTGGACCTGACAGACGCTGAACCTGCGACCTGGTGGTTATG 3300
Db 2129 GAATGCTGAAGATTTTGGACCTGACAGACGCTGAACCTGCGACCTGGTGGTTATG 2188
QY 3301 ATTATACACTGTTGAACTATCTGTGACAGTGTGACAGTCTCAGCAGATGAAACCCGTG 3360
Db 2189 ATTATACACTGTTGAACTATCTGTGACAGTGTGACAGTCTCAGCAGATGAAACCCGTG 2248
QY 3361 CTGTGTGAGAGCACTCTGAGAGAGTCTGTCTATCTATCTGATTGTTGTTTCATCCAGAA 3420
Db 2249 CTGTGTGAGAGCACTCTGAGAGAGTCTGTCTATCTGATTGTTGTTTCATCCAGAA 2308
QY 3421 ACAATGCTACATGATGTGAGAACTTACACAAAGATAGAGATTTTCTGTGCTCAGAGCAG 3480
Db 2309 ACAATGCTACATGATGTGAGAACTTACACAAAGATAGAGATTTTCTGTGCTCAGAGCAG 2368
QY 3481 GGTGAAATCACTGAAGGCTCTGTTCTTTCATCATTA 3518
Db 2369 GGTGAAATCACTGAAGGCTCTGTTCTTTCATCATTA 2406

```

RESULT 6

ADJ38208 standard; DNA; 2406 BP.

ADJ38208;

06-MAY-2004 (first entry)

Plastid division-related Arc6 orthologue gene 2.

XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
 XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
 XX herbicide target; gene; ds.

Arabidopsis thaliana.

W02004001003-A2.

31-DEC-2003.

20-JUN-2003; 2003WO-US019536.

XX 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 XX (UNMS) UNIV MICHIGAN STATE.
 PI Oseveryoung KW, Vilha S, Koksharova OA, Gao H;
 XX WPI; 2004-082486/08.
 DR P-PSDB; ADJ38207.
 DR
 PT New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287pp; English.

CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a gene which is
 CC related to the invention.

SQ Sequence 2406 BP; 612 A; 493 C; 618 G; 683 T; 0 U; 0 Other;

Query Match 46.8%; Score 1717.6; DB 12; Length 2406;
 Best Local Similarity 79.1%; Pred. No. 0;
 Matches 2402; Conservative 0; Mismatches 4; Indels 632; Gaps 5;

```

QY 481 ATGAGACCTGAGTCACTGCGGATTTGTTCTCCCATTCGAATTATGCGATTACCA 540
Db 1 ATGAGACCTGAGTCACTGCGGATTTGTTCTCCCATTCGAATTATGCGATTACCA 60
QY 541 CCGCGACGACAAAGCTCCGACGACGACCAACACCTCTTACACTATCTGTCGCGCAGC 600
Db 61 CCGCGACGACAAAGCTCCGACGACGACCAACACCTCTTACACTATCTGTCGCGCAGC 120
QY 601 AAATGGGCGACGCTTCTCTCCGACTTCAATTACCTCCGATTCCTCTCTCTCTCC 660
Db 121 AAATGGGCGACGCTTCTCTCCGACTTCAATTACCTCCGATTCCTCTCTCTCTCC 180
QY 661 TTCGCAACGCGCACACCAACGCGCACTCTGCTCTCTGCGCAACATCTATTGATCGTCC 720
Db 181 TTCGCAACGCGCACACCAACGCGCACTCTGCTCTCTGCGCAACATCTATTGATCGTCC 240
QY 721 GAACGCCACGTCGCCATCCCATTTGATTCTTACCAAGGATTTAGAGCTCAAAACATTTTC 780
Db 241 GAACGCCACGTCGCCATCCCATTTGATTCTTACCAAGGATTTAGAGCTCAAAACATTTTC 300
QY 781 TTAAACGATGAAATCAAGAGCAATTGAGGCTTGTGAAACCGCGCAATTGCGGT 840
Db 301 TTAAACGATGAAATCAAGAGCAATTGAGGCTTGTGAAACCGCGCAATTGCGGT 360
QY 841 TTCAGCAGCAGCGCTTTAATACGCGGAGACAGATTCTTCAAGCTGCTTGGAAACTCTG 900
Db 361 TTCAGCAGCAGCGCTTTAATACGCGGAGACAGATTCTTCAAGCTGCTTGGAAACTCTG 420
QY 901 TCTAATCTCGGTCTTAAAGAGATACAGAGAGTCTTCTTGAATGATGAAGAGCTACA 960
Db 421 TCTAATCTCGGTCTTAAAGAGATACAGAGAGTCTTCTTGAATGATGAAGAGCTACA 480
QY 961 GTCATCATGATGTTCTTGGGATTAAGTAATTGATTCGAAATATAAGTTCTTC 1020
Db 481 GTCATCATGATGTTCTTGGGAT----- 504
QY 1021 GTTTAATTTCAATGAATGATTAAGAGAGAACTTTATCTAGTGAAGTTCTCGGCGC 1080

```

D	b		505	-	-----AAGTTCCTGGGCGC	518
O	y		1081	TCTCTGTATTTGCAGAAGAGTGATGACCTGAGTAAGTCTTTCGGGTGGTAGAGCCCT	1140	
D	b		519	TCTCTGTGTATTGCCAAGAAAGGTGTGAGACTGAGATAGTCTTTCGGGTGGTAGAGCCCT	578	
O	y		1141	GCTTAAAGAGAGGTGGCCTTAAGTCCCTTAAAGCAATAATGTGGTTTTAGTTAATGGGCTTTC	1200	
D	b		579	GCTTAAAGAGAGAGGTGGCCTTAAGTCCCTTAAAGCAATAATGTGGTTTTAGTTAATGGGCTTTC	638	
O	y		1201	GTTTCTCGATGTCTCGAGGAGATGCTATAGGCAATTGGATCACCTGAATTTTATTAATTCGTGTA	1260	
D	b		639	GTTTCTCGATGTCTCGAGGAGATGCTATAGGCAATTGGATCACCTGAATTTTATTAATTCGTGTA	698	
O	y		1261	TGAGTTTGTGAGGAAGCTTTGAAAGCTTTTACAAGTAAGTGAATGACTTGTCTGTGAATTTTG	1320	
D	b		699	TGAGTTTGTGAGGAAGCTTTGAAAGCTTTTACAAGTAAGTGAATGACTTGTCTGTGAATTTTG	730	
O	y		1321	ACGAGCGTTGGCTTTATPAAGAACCTTCTTGATTTGATACCTTGTGTTATTTGACTTGTGTGA	1380	
D	b		731	-	-----A	731
O	y		1381	GGAGGAAGAGCAAGTAGCCTTGCACCGGATTTATAGTGCACAAATTGATGACCTTTGGA	1440	
D	b		732	GGAGGAAGAGCAAGTAGCCTTGCACCGGATTTATAGTGCACAAATTGATGACCTTTGGA	791	
O	y		1441	AGAGATCACTCCGCGTATATGTCCTTGAGACTATCTTGGCTTACCGCTGTGTGATGATTAAGC	1500	
D	b		792	AGAGATCACTCCGCGTATATGTCCTTGAGACTATCTTGGCTTACCGCTGTGTGATGATTAAGC	851	
O	y		1501	TGCGAAAAGACTAAATGGTTTAAGCGGTGTGCGAAATTTTGTGTCTGTGAGAGAG	1560	
D	b		852	TGCGAAAAGACTAAATGGTTTAAGCGGTGTGCGAAATTTTGTGTCTGTGAGAGAG	911	
O	y		1561	TGGAACATCAAGCTCTTGTGTGGGGGTTTGACCCTGTGAGAGTTTAAATGAGGGGTTTTT	1620	
D	b		912	TGGAACATCAAGCTCTTGTGTGGGGGTTTGACCCTGTGAGAGTTTAAATGAGGGGTTTTT	971	
O	y		1621	ACGAATGACAGCTGCTGAGCAAGGTATACAGTTTAAATCTTTTAAATTTCTTAAACA	1680	
D	b		972	ACGAATGACAGCTGCTGAGC-----	991	
O	y		1681	TGATATPAACTTAAAGTTCATTTTAAATGATATGTGTGTGATGATCTTTTGTGA	1740	
D	b		992	-	-----AAGTTGACTTTTGTGA	1008
O	y		1741	GCTACCCCAGCAAGATATATCCAGCAGATCATTTGAAAGTTTGAAGTTGCACTTGCTCTT	1800	
D	b		1009	GCTACCCCAGCAAGATATATCCAGCAGATCATTTGAAAGTTTGAAGTTTGAAGTTGCACTTGCTCTT	1068	
O	y		1801	GTTGCTCAAGCTTTTATTTGTATGAAGCAACACTTTTACAGSANTGCTGATPAACAAATTC	1860	
D	b		1069	GTTGCTCAAGCTTTTATTTGTATGAAGCAACACTTTTACAGSANTGCTGATPAACAAATTC	1128	
O	y		1861	CAGCAACTTCAGCAGGCTTAAGGTAAATGCTATGAGATTCCTGCGATGTTGATGATACA	1920	
D	b		1129	CAGCAACTTCAGCAGGCTTAAGGTAAATGCTATGAGATTCCTGCGATGTTGATGATACA	1188	
O	y		1921	CGGAATTAATTGGGAATATGACTTTCGCTCTTAAAGAGGGAATCTGTGCACTCTTATAGGC	1980	
D	b		1189	CGGAATTAATTGGGAATATGACTTTCGCTCTTAAAGAGGGAATCTGTGCACTCTTATAGGC	1248	
O	y		1981	AAGATTGATGAATGCGCTATGTGTGTTGGGCTTACAGACTGAGAGATTCACATPATAGAAAT	2040	
D	b		1249	AAGATTGATGAATGCGCTATGTGTGTTGGGCTTACAGACTGAGAGATTCACATPATAGAAAT	1308	
O	y		2041	CCAAGTATTTGAGACTTTGTTTGGAGATTCAAATTCGTATGACAAATGATGATCTCCCT	2100	
D	b		1309	CCAAGTATTTGAGACTTTGTTTGGAGATTCAAATTCGTATGACAAATGATGATCTCCCT	1368	
O	y		2101	GGAATATGCAATTTGTGGAAACCTGTGTTGGAGGGGTTGTCTTCTTAAGSTTCAGAGAC	2160	
D	b		1369	GGAATATGCAATTTGTGGAAACCTGTGTTGGAGGGGTTGTCTTCTTAAGSTTCAGAGAC	1428	

QY	2161	ACCAAGATAAAAATTAAAACTCGGGGACTACTATGATATCTATGTTTGAAGTTAC	2220
DB	1429	ACCAAGATAAAAATTAAAACTCGGGGACTACTATGATATCTATGTTTGAAGTTAC	1488
QY	2221	TTGGAAGAAGTGAAGTAGTTCAAGGTTCTTCCTTTAGCTCTGCTGCACTATATGCAAGG	2280
DB	1489	TTGGAAGAAGTGAAGTAGTTCAAGGTTCTTCCTTTAGCTCTGCTGCACTATATGCAAGG	1548
QY	2281	ATTGAGCCGAGCATGTGAAGCTATGTCTATGCAAGGCACTGAGAAAGTTTTTCTCTCC	2340
DB	1549	ATTGAGCCGAGCATGTGAAGCTATGTCTATGCAAGGCACTGAGAAAGTTTTTCTCTCC	1608
QY	2341	CGCTATACAGTAAATCTGGGCTGAACCCAGAGATGTGCAAGAGACAGTGTTAGTGTA	2400
DB	1609	CGCTATACAGTAAATCTGGGCTGAACCCAGAGATGTGCAAGAGACAGTGTTAGTGTA	1668
QY	2401	GATCCTGTGTGTAACTATGTAGCCGTGATGTGTGAGCCCTGATGCTTTATTTGCAAGAGCT	2460
DB	1669	GATCCTGTGTGTAACTATGTAGCCGTGATGTGTGAGCCCTGATGCTTTATTTGCAAGAGCT	1728
QY	2461	GTAAGACCCCTTGAAAACTTTGAAACTATATGATTATGCAATTGAGGTGGGGTCTCAGAG	2520
DB	1729	GTAAGACCCCTTGAAAACTTTGAAACTATATGATTATGCAATTGAGGTGGGGTCTCAGAG	1788
QY	2521	AGTAGCGTGTATGAATCTACTGTGAATATCCGTGCTGATATGTATGAAGAGGCAAGT	2580
DB	1789	AGTAGCGTGTATGAATCTACTGTGAATATCCGTGCTGATATGTATGAAGAGGCAAGT	1848
QY	2581	GTGAAGATCTAGTGTCTGTGTGTGCAATTGAGCTGATTTCACTGTTCACGCCAAGAT	2640
DB	1849	GTGAAGATCTAGTGTCTGTGTGTGCAATTGAGCTGATTTCACTGTTCACGCCAAGAT	1908
QY	2641	TTTCTTAAAAAGCAGCTCATCTTTTCAAGGCAAGGATATGTGTTCTTCTATGAATCGAT	2700
DB	1909	TTTCTTAAAAAGCAGCTCATCTTTTCAAGGCAAGGATATGTGTTCTTCTATGAATCGAT	1968
QY	2701	GTGCTACCAATAGGATATGATTAATATGATGCAATTTTCATATATCTGCATTTCTCAAAATA	2760
DB	1969	GTGCTACCA-----	1978
QY	2761	TGCTTGTTTTGTAGCTATAGAACATATGTTCCACTTAATATCATATGTCCCAAAAGTTGTACC	2820
DB	1979	-----	1978
QY	2821	AAGATTAAACAAGTGTGTAGTAATATTTCACTAAATTATGTCTGTGAATTTTTTGATCAAA	2880
DB	1979	-----	1978
QY	2881	CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTGAATATACGTAGAGATTAG	2940
DB	1979	-----	1978
QY	2941	AGATTGCTTAGTGTGTGCTTTGTCCAACTTTTCTTTCTGATTATTTTCTTTTCATTT	3000
DB	1979	-----	1979
QY	3001	AGGTCAGTCAGACTGATCAATTCAGAGCACTTCCAGATGAGATCTAGACTGACA	3060
DB	1980	AGGTCAGTCAGACTGATCAATTCAGAGCACTTCCAGATGAGATCTAGACTGACA	2039
QY	3061	GAAATATAGTATCCAGTGGCAGAAAGATTAAAGTCTGTGGCTTTGGGCTGTATGACGCGAT	3120
DB	2040	GAAATATAGTATCCAGTGGCAGAAAGATTAAAGTCTGTGGCTTTGGGCTGTATGACGCGAT	2099
QY	3121	AGAAATGTATCCAGAGGTGAGGGAATTAATCTACAATTCATCAATGATTTGTGAAAACTGT	3180
DB	2100	AGAAATGTATCCAG-----	2113
QY	3181	TGAGCATGATTAATATGTCTGTGCTTGTGTTGATTTCTGTATTTATATAGTTTGTGATGGC	3240
DB	2114	-----AGTTTGTGATTTGGGC	2128

QY 861 CAGCCGAGACAGATTTCTTCAAGCTGCTGCGAACTGCTTAATCCCTGCTTGAAG 920
 DB 311 TAGTGTGTGATGATTTCTTCAAGCTGCTGCTTGAAGCCCTAGCTATCTGCTTGAAG 370
 QY 921 AGAGTACATGAGAGTCTTCTTGAATGATGAGAGCTACATCATCATGATGCTTGG 980
 DB 371 AGAGTATATCAAGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 405
 QY 981 GATATAGATATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 1040
 DB 406 -----AAGATGAGAGATCTT 420
 QY 1041 ATAAAGAGAGAGATTTATCTAGTGAAGTCTGCGGCTCTGCTGATTTGAAGAAG 1100
 DB 421 CCATTCTCACTGAATCCCTTTCGACAAAGTCTGCTGAGCTGCTGCTGCTGCTGCTG 480
 QY 1101 GTGTGAGACATGAGATGCTTCTGCGGCTGCTGAGGCTCTGCTTGAAGAGAGCTTGA 1160
 DB 481 CTGAGAGAGAGAGAGTGTGCTTCCGATTTGAGAGGCTTACTGAGAGAGAGAGAGAG 540
 QY 1161 AGTGTATTAAGCAAGATGCTGCTTGAATGAGAGGCTGCTTCTGCTGATTTGAGAGG 1220
 DB 541 AGATGTTTAAGCAAGATGCTGCTTGAATGAGAGGCTGCTTCTGCTGATTTGAGAGG 600
 QY 1221 ATGCTATGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
 DB 601 ATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 652

RESULT 10

ID ADJ38206 standard; DNA; 2283 BP.
 AC ADJ38206;

06-MAY-2004 (first entry)

DE Plastid division-related Arc6 orthologue gene 1.

XX prokaryotic type; plastid division; Ftn2; AR6; ARCS; Fzo; plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KM herbicide target; gene; ds.

OS Oryza sativa.

PN WO2004001003-A2.

PD 31-DEC-2003.

PE 20-JUN-2003; 2003WO-US019536.

PR 20-JUN-2002; 2002US-0390140P.

PR 09-AUG-2002; 2002US-0402242P.

PR 20-JUN-2003; 2003US-00600070.

PA (UNMS) UNIV MICHIGAN STATE.

P1 Oosteryoung KM, Vitha S, Koksharova OA, Gao H;

DR MPI; 2004-082486/08.

DR P-PSDB; ADJ38205.

PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.

PS Disclosure; Fig 8; 287pp; English.

CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (AR6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further

CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a gene which is
 CC related to the invention.

SQ Sequence 2283 BP; 551 A; 576 C; 592 G; 564 T; 0 U; 0 Other;

Query Match 5.3%; Score 194; DB 12; Length 2283;

Best Local Similarity 55.6%; Pred. No. 2.1e-39; Mismatches 300; Indels 110; Gaps 3;

Matches 514; Conservative 0; Mismatches 300; Indels 110; Gaps 3;

QY 1368 TGAGCTTGTGTAG 1427
 DB 716 TCAAGCTTTCAG 775
 QY 1428 ATGAGACTTTGAG 1487
 DB 776 ATGAAACTCTGAG 835
 QY 1488 GTGATGATTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1547
 DB 836 ACACAG 895
 QY 1548 CTGTTGAG 1607
 DB 896 GCGTTGAG 955
 QY 1608 ATGAG 1667
 DB 956 ACGAG 993
 QY 1668 AATTTCTTACAG 1727
 DB 994 ----- 993
 QY 1728 TGATCTTTTGTAG 1787
 DB 994 -GATTTCTTTTCAAAAACAG 1052
 QY 1788 TGACCTTGTCTTGTGCTCAAGCTTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1847
 DB 1053 AGCACTTGCATGATGCTGCTCAAGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1112
 QY 1848 TGATGAG 1907
 DB 1113 GATGATCTTTTGTGAG 1164
 QY 1908 GTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1967
 DB 1165 -----TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1208
 QY 1968 ACTGCTTATAG 2027
 DB 1209 ATTGCTATGAG 1268
 QY 2028 ACAATATGAG 2084
 DB 1269 ACCATACAG 1328
 QY 2085 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2144
 DB 1329 GAT 1388
 QY 2145 TCGTAGTTCAG 2204
 DB 1389 TCGTAGGAG 1448
 QY 2205 TAGGTTTGAAGTACTTGAAG 2264
 DB 1449 AGAAGTTTGAAGTACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1508
 QY 2265 TGCAACTATGAG 2288

Db 1509 TGTGCTATTGCAAAACTTGCTGC 1532

RESULT 11

ACN48781
ID ACN48781 standard; cDNA; 607 BP.

AC ACN48781;

DT 02-DEC-2004 (first entry)

DE Cotton primed seed EST Clone ID: LIB3825-027-06-Kc-H1, SEQ:3562.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
XX variety DP50B; library LIB3825; molecular tag; molecular marker;
KM genetic mapping; molecular mapping; seed germination; plant growth;
XX plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

XX US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.

XX (FENG/) FENG P C C.

XX (FINC/) FINCHER K L.

XX (ZIEG/) ZIEGLER T E.

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX WPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its
XX fragment, useful for isolating a variety of agronomically significant
XX genes associated with plant growth, quality or yield, and as molecular
XX tags to map genes.

XX Claim 1; SEQ ID NO 3562; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
XX ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
XX from primed or non-primed seeds from variety DP50B, mature seeds from
XX variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
XX tissue, developing fibres, carpel walls and septa from variety
XX Nicotian3B. The invention also relates to substantially purified
XX proteins or their fragments encoded by nucleic acid molecules of the
XX invention, and to transformed plants having a nucleic acid construct
XX comprising a nucleic acid of the invention. The cotton ESTs are useful as
XX molecular tags to isolate genetic regions, to isolate genes, to map
XX genes, to determine gene function and to determining whether genes are
XX members of a particular gene family. The nucleic acid molecules may be
XX used for isolating a variety of agronomically significant genes
XX associated with plant growth, quality, yield, and could also serve as
XX links in metabolic and catabolic pathways. The nucleic acid molecules are
XX also useful for identifying genes important in initiating and maintaining
XX seed germination or that may be used to mitigate stresses encountered
XX during seed germination. The ESTs additionally enable the acquisition of
XX promoters and cis-regulatory elements which will be useful to express
XX agronomically significant genes in these tissues and/or other tissues,
XX and also permits the acquisition of molecular markers useful in breeding
XX schemes, genetic and molecular mapping, and in cloning of agronomically
XX significant genes. The nucleic acid molecules are further useful for
XX detecting the expression level or pattern of a protein or mRNA and for
XX detecting the presence or quantity of a protein by tissue printing. The
XX present sequence represents a specifically claimed EST isolated from a
XX cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
XX data for this patent did not form part of the printed specification, but

CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?docID=US20040123340
XX
XX Sequence 607 BP; 172 A; 103 C; 164 G; 167 T; 0 U; 1 Other;

Query Match 4.9%; Score 178.6; DB 13; Length 607;
Best Local Similarity 69.3%; Pred. No. 1.1e-35;
Matches 266; Conservative 0; Mismatches 100; Indels 18; Gaps 1;

Qy 1986 TGAATGATGCCGATATGATGTTGGCTTACAGACATGAGATTCACAAATATAGAAATCCAGC 2045

Db 5 TGNAGAGTGCCTGTCGTCGTTGGCTTACAGACATGATAGCTCCCTTATGAAATATCATC 64

Qy 2046 TATTGTGAGTGTGTTTGGAGAAATTCAAATCGATGACAAATGATGATCTCCCTGACT 2105

Db 65 TATTGTGAAATTTGTTGTTGAAAATCAAGATGACATGACAGAGATCTTCCGGGCT 124

Qy 2106 ATGCAAATTTGTGAAAACCTGTGTGGACAGGGGTGTCCTTCCATGATTCAGACACCA 2165

Db 125 TTGCMAACTGTCTGAGGATGCTTAATGAGGTGTTTTCATGATTTAGACACCAA 184

Qy 2166 AGATTAATAATTTAACTCGGGACTACTATGATGATCCTATGATTTTGGATTACTTGA 2225

Db 185 AGATATACAAATTCAGCTTGGAGATTTATATGATGATCCTACTGTCCTGAGATTTTGA 244

Qy 2226 AAGATGAGAGTATGTTCAAGGTTCTCCTTAACTGCTGCAACTATGCAAGATTTGG 2285

Db 245 AAGCTTGAGAGGAGCAGAGTGTTCACCTTGCTGCGACCGCAGCTATATGAGATTTAG 304

Qy 2286 AGCCGAG-----CATGGAAGCTAGTCTATGCAAGCACTGCAGAA 2327

Db 305 TGCTGAGGCTACTGACGTTCTTGATCATGTAAAGGCTATGCAATTCAGGATTTGAGAA 364

Qy 2328 AGTTTTCCTTCCCGCTATACAGA 2351

Db 365 GGTGTTTCTCTTCTGTCGCTCAGA 388

RESULT 12

ADP93143

ID ADP93143 standard; cDNA; 439 BP.

XX ADP93143;

XX 09-SEP-2004 (first entry)

XX Cotton expressed sequence tag, EST, #2154.

XX Cotton; ss; EST; expressed sequence tag; plant; plant protection;
XX plant improvement; marker-assisted breeding.

XX Gossypium hirsutum; variety Nuocotton33B.

XX US2004123338-A1.

XX 24-JUN-2004.

XX 08-DEC-2000; 2000US-00732627.

XX 10-DEC-1999; 99US-0170255P.

XX (FINC/) FINCHER K L.

XX Fincher KL;

XX WPI; 2004-479807/45.

XX New substantially purified nucleic acid molecule that encodes a cotton
XX protein or its fragment, useful as molecular tool for the targeting and
XX isolation of novel genes for plant protection and improvement.

XX Claim 1; SEQ ID NO 2154; 30pp; English.

CC The invention relates to a substantially purified nucleic acid molecule
CC that encodes a cotton protein or its fragment comprising an EST
CC (expressed sequence tag) appearing as ADB90990-AD95919. Also included
CC are a substantially purified cotton protein or its fragment encoded by a
CC nucleic acid molecule above and a transformed plant (having a nucleic
CC acid molecule which comprises: an exogenous promoter region which
CC functions in a plant cell to cause the production of a mRNA molecule; a
CC structural nucleic acid molecule comprising one of the ESTs or their
CC complements; a 3' non-translated sequence that functions in the plant
CC cell to cause termination of transcription and addition of polyadenylated
CC ribonucleotides to a 3' end of the mRNA molecule). The ESTs are useful as
CC a molecular tool for the targeting and isolation of novel genes for plant
CC protection and improvement. The ESTs are useful for developing new
CC strategies for understanding critical plant developmental and metabolic
CC pathways, for isolating genes and promoters, for identifying and mapping
CC the genes involved in developmental and metabolic pathways, and for
CC determining gene function. The cotton nucleic acid molecules are useful
CC as molecular tags to isolate genetic regions, isolate genes, map genes,
CC and determine gene function. The nucleic acid molecules are useful for
CC determining if genes are members of a particular gene family and for use
CC in marker-assisted breeding programs. The present sequence is one of the
CC 4930 cotton ESTs of the invention. NOTE: The sequences are not displayed
CC in the specification but are available in electronic format from the
CC USPTO at seqdata.uspto.gov/sequence.html?DocId=20040123338.

SQ Sequence 439 BP; 87 A; 153 C; 79 G; 119 T; 0 U; 1 Other;

Query Match	Score	DB	Length
4.8%	175	12	439

Best Local Similarity 65.0%; Pred. No. 8.1e-35;

Matches 278; Conservative 0; Mismatches 141; Indels 9; Gaps 1.

Qy	499	GTGGGCAATGGTCTCTCCCAATTCCAATTATGCGATTAACACCGGCGGACGACAAAGTC	558
Db	12	GTTAGCACTGGGCTGTGCACTCCGACATTGTCTCTTATTGCACTCTGTAAACCTTCC	71
Qy	559	CGAGCTAGGCCAACACCTCTACAACTATCTGCTCCGCGAGCAAAATGGCGACCGCTTT	618
Db	72	ANACTCCACACCGGTTCAACTACCGCTCGTGTGCTCCGCGACGAAATATGGGCGAGCGCTTC	131
Qy	619	CTCTCCGACTCAATTTCACTCCG-----ATTCTCGTCCCTCCCTTGGCGAC	669
Db	132	CTGTGTACTTCCAAATTTCTCCCTGCGCCCGCATTAATCCGTCTCTCTTCTCTCTCTCA	191
Qy	670	GCGACACACACCGGCACCTCTGCTCTCTGCGCCACCATCTATTGATGTCGCCGACGCCAC	729
Db	192	ACCGGCACTTTTCTCCCCCTTAACCTCCCTCCCTCTGCGCCCTTCTCTCCGAAAGCCAC	251
Qy	730	GTCCCATCCCCATGATTTCTTACCAAGGATTAGAGACTCAAAACATTTCTTAAACGAT	789
Db	252	GTTCCTCATCCCTCTGATTTCTTACAGGTTTATAGAGCGGACACTCATTTCTTAGGGTAT	311
Qy	790	GGATTCAGAAAGAGCTGGAAGCTAGGGTTTGAAACCGCGGCATTTGGGTTTCAGCGAC	849
Db	312	GGAAATTAGAAAGAGCTTATAGAACGAGGTTTGAAGACGGCTCTCAATATATGGTTCAAGTCAA	371
Qy	850	GACGCTTTAATCAGCGCGAGACAGATTTCTTCAAGCTGCTTGGAAACTCTGTCTAATCT	909
Db	372	GACACCAATTAATAGCGGAAAGACAGATTTCTTCTAGCTGCTGTGAACCCCTATCAACCT	431
Qy	910	CGGCTTAG 917	
Db	432	GGCTCTTAG 439	

RESULT 13

ID ADJ38255 standard; cDNA; 545 BP.

AC ADJ38255;

DT 06-MAY-2004 (first entry)

Plastid division-related Arc6 orthologue cDNA 42.

XX
KM prokaryotic type; plastid division; Ftn2, ARC6; ARC5; Fzo; plant cell
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KM herbicide target; gene, 88.

05 *Gossypium arboreum.*

PN WO2004001003-A2

PD 31-DEC-2003

20-JUN-2003; 2003WO-US019536.

20-JUN-2002; 2002US-0390140P.

PR 20-JUN-2003: 2003US-00600070.

AA
PA
(TIMMS) INTV MICHIGAN STATE-

XX		
PI	Observing KW	Vitha S Koh

XX
WB1: 2004-082495/008

11-37

PT New isolated Fru2, AHC5 and/or Fru2-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.

PS Disclosure; Fig 8; 287pp; English.

This invention relates to novel prokaryotic type or plasmid division and related genes and proteins. In particular, the invention relates to novel *fnr2* (*ARC6*), *ARC5* and *Pro*-like genes and polypeptides. The methods and compositions of the present invention are useful for further characterizing plasmid division in plant cells, in order to vary agronomic and horticultural characteristics of economically important plants, such as crop, ornamental and woody plants. They can also be used as herbicide targets. The present sequence is a cDNA sequence which is related to the invention.

SQ Sequence 545 BP; 164 A; 107 C; 120 G; 151 T; 0 U; 3 Other;

Query Match 4.4%; Score 161; DB 12; Length 545;

Best Local Similarity 69.6%; Pred. No. 4e-31;

Matches 218; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

3224 ATAGGTTTGGATGGGCGAATGCTGAAGATTGGACTGACAGAGCAGCTGAACTGCCA 3283

Db 82 AGAGTTCGATGTC AATGTTGAAGACATGCAGATCGTGCAGCCGAATCGCTCA 141

3284 GCTTGGGTTGTTATGATTATACACTGTTGAACCTATCTGTTGACAGTGACAGTCTC 3343

Db 142 GCTTGGTTGGGTATGATATAGTCTACTGACATGGCCATTGACAGTGTACCCCTTC 201

3344 AGCAGATGGAACCCGCTCTGCTGGAAGCAACTCTGGAGGAGTCTGCTTGTCTATCTGA 3403

202 ACTAGATGCCAGCGAGCTGTACTCGAAGCTACTCTGGAAGATCCACCCTGCTTGACTGA 261

3404 TTTTGGTTTCAATCCAGAAACAATGCTACTTGATGTCAGAACCTTACACACAAAGATACGAAGT 3463

323 321

3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542 3543 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556 3557 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570 3571 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584 3585 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598 3599 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612 3613 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626 3627 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640 3641 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654 3655 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668 3669 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682 3683 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696 3697 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710 3711 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724 3725 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738 3739 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752 3753 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766 3767 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780 3781 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794 3795 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808 3809 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822 3823 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836 3837 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850 3851 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864 3865 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878 3879 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892 3893 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906 3907 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920 3921 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934 3935 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948 3949 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962 3963 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976 3977 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990 3991 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004 4005 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018 4019 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032 4033 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046 4047 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060 4061 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074 4075 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088 4089 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102 4103 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116 4117 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130 4131 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144 4145 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158 4159 4160 4161 4162 4163 4164 4165 4166 4167 4168 4169 4170 4171 4172 4173 4174 4175 4176 4177 4178 4179 4180 4181 4182 4183 4184 4185 4186 4187 4188 4189 4190 4191 4192 4193 4194 4195 4196 4197 4198 4199 4200 4201 4202 4203 4204 4205 4206 4207 4208 4209 4210 4211 4212 4213 4214 4215 4216 4217 4218 4219 4220 4221 4222 4223 4224 4225 4226 4227 4228 4229 4230 4231 4232 4233 4234 4235 4236 4237 4238 4239 4240 4241 4242 4243 4244 4245 4246 4247 4248 4249 4250 4251 4252 4253 4254 4255 4256 4257 4258 4259 4260 4261 4262 4263 4264 4265 4266 4267 4268 4269 4270 4271 4272 4273 4274 4275 4276 4277 4278 4279 4280 4281 4282 4283 4284 4285 4286 4287 4288 4289 4290 4291 4292 4293 4294 4295 4296 4297 4298 4299 4300 4301 4302 4303 4304 4305 4306 4307 4308 4309 4310 4311 4312 4313 4314 4315 4316 4317 4318 4319 4320 4321 4322 4323 4324 4325 4326 4327 4328 4329 4330 4331 4332 4333 4334 4335 4336 4337 4338 4339 4340 4341 4342 4343 4344 4345 4346 4347 4348 4349 4350

[illegible]

0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99

RESULT 14

ACN48855/c

ID ACN48855 standard; cDNA; 552 BP.
 XX ACN48855;
 XX
 XX 02-DEC-2004 (first entry)
 DT
 XX Cotton primed seed EST Clone ID: LIB3825-027-06-N6-H1, SEQ3636.
 DE
 XX Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
 KM variety DP50B; library LIB3825; molecular tag; molecular marker;
 KM genetic mapping; molecular mapping; seed germination; plant growth;
 KM plant quality; plant yield; plant breeding; tissue printing; ss.
 OS
 XX *Gossypium hirsutum*.
 XX
 XX US2004123340-A1.
 XX
 XX 24-JUN-2004.
 PD
 XX 12-DEC-2001; 2001US-00021323.
 PF
 XX 14-DEC-2000; 2000US-0255619P.
 PR
 XX (DEIK/) DEIKMAN J.
 PA (FENG/) FENG P C C.
 PA (FINC/) FINCHER K L.
 PA (ZIEG/) ZIEGLER T E.
 XX
 PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;
 PI WPI; 2004-479808/45.
 DR
 XX
 XX
 PT New isolated nucleic acid molecule that encodes a plant protein or its
 PT fragment, useful for isolating a variety of agronomically significant
 PT genes associated with plant growth, quality or yield, and as molecular
 PT tags to map genes.
 PT
 XX
 XX
 PS Claim 1; SEQ ID NO 3636; 34pp; English.
 CC
 CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN43220-ACN63099). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety DP50B, mature seeds from
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
 CC tissue, developing fibres, carpel walls and septa from variety
 CC Nicotiana3B. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
 CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determine whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be
 CC used for isolating a variety of agronomically significant genes
 CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining
 CC seed germination or that may be used to mitigate stresses encountered
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes, genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from the US patent office at
 CC seqdata.uspo.gov/sequence.html?pcid=US20040123340
 CC
 XX
 XX Sequence 552 BP; 153 A; 122 C; 110 G; 167 T; 0 U; 0 Other;
 Query Match 4.3%; Score 159.4; DB 13; Length 552;

Best Local Similarity 69.3%; Pred. No. 1.1e-30;
 Matches 217; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
 QY 3224 ATAGTTTGGATGGCGGAATGCTGAAGATTGGACTGACAGACAGCTGAAACTGGCA 3283
 DB 444 AGAGGTTTCGATGCTGAATGTTGAAGACATGGAACAGATCGAGCCGAAATCGCTCA 385
 QY 3284 GCTTGGGTTGTTTANGATTATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTC 3343
 DB 384 GCTTGGTGGTATATGAATATAGTCTACAGAACTAGCCATTTGACAGTGTACCCCTTC 325
 QY 3344 AGCAGATGGAACCCGCTCTGGTGGAGGAACTGAGAGCTGAGAGCTGTGTATCTGA 3403
 DB 324 ACTAGATGGCCAGCGAGCTATAGTCGAGTACTCTGGAAGAAATTCACCTGCTTGA 265
 QY 3404 TTTGGTTTCATCCAGAAAACAACTGCTATGATGTACAGAACTTACACAAAGATACGAGT 3463
 DB 264 TGTTCATCATCCGAGAACAACTGCTTATGTATTAATCCTTACACAGAGATATGAT 205
 QY 3464 TTTCTGTCCAGTCCAGGTTGAAAATCACTGAAAGCTGTGTTCTTGATCATATATAC 3523
 DB 204 GTCTTGTTCCACTCAGGCTGAAAAATCACTGAAGATCTGTACAAATCTTAGTATG 145
 QY 3524 TCATATGTAGCAT 3536
 DB 144 ATGTATTAAGCAT 132
 RESULT 15
 ADJ38217
 ID ADJ38217 standard; cDNA; 608 BP.
 XX
 XX ADJ38217;
 AC
 XX 06-MAY-2004 (first entry)
 DT
 XX
 XX Plastid division-related Arce.orthologue cDNA 10.
 DE
 XX prokaryotic type; plastid division; Ftn2; ARCE; ARCS; Fzo; plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KM herbicide target; gene; ss.
 XX
 XX Solanum tuberosum.
 OS
 XX WO2004001003-A2.
 PN
 XX 31-DEC-2003.
 PD
 XX 20-JUN-2003; 2003WO-US019536.
 PF
 XX 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 XX (UNMS) UNIV MICHIGAN STATE.
 PA
 XX OeserYoung KW, Vitsha S, Koksharova OA, Gao H;
 PI WPI; 2004-082486/08.
 DR
 XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 PT
 XX
 XX Disclosure, Fig 8; 287bp; English.
 ES
 XX This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important

CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.

XX
SQ Sequence 608 BP; 146 A; 167 C; 128 G; 167 T; 0 U; 0 Other;

Query Match 4.3%; Score 158; DB 12; Length 608;

Best Local Similarity 63.2%; Pred. No. 2.6e-30;

Matches 264; Conservative 0; Mismatches 145; Indels 9; Gaps 1;

```
QY 581 CAACTATCTGCTCCGACGAAATGGGCCGACCGCTTCTCTCGACTTCAATT----- 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 CTACTAATCTCTCCGCTAGTAATGGCGGATCGCTTCTCGCGATTTCCAATTCCTTC 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 636 ----CACCTCGATTCCTCCCTCCCTCTCTCTCGCACCGCACCGCCCACTCTCG 691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 CTTCCACCCACACCTCCGACTCATCGATTCAGAAATTCAACTTCAACAACCTCCGTTA 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 692 TCTCTGCGACCATCTATATGATGTCGCCGAACGCAAGTCCCATCCCATTTGATTCT 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 CGACTATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 752 ACCAGTATTAGAGCTCAACACATTTCTTAACCGATGAATCAAGAGACATTGAA 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 ATAGAGTGTGTGTGTAAGCTCACTTCTCTGTAAGGTATAGAGATGTACGATG 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 812 CTAGGGTTTGAACCCGCCGAATTCCGTTTCAAGCAGAGCTTTAATCAGCCGAGAC 871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 CTAGAAATTACAAAGCTCCGAGTAGAGATACAGTCAAGAGACATTTGATGGCCGAC 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 872 AGATTCTTCAAGCTGTGCGAACTGTCTAATCTCGCTAGAAAGAGTACAATG 931
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 AGATTCTTCAAGCTGTGCGAACTGTCTAATCTCGCTAGAAAGAGTACAATG 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 932 AAGGCTCTTGTGATGATGAAGAGCTACAGTCATCACTGATGTTCTTGGGATTAAGT 989
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 AAGGCTCTGCTCAGCATGATGATGATTAATTAATTAATTAATTAATTAATTAAT 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: June 9, 2005, 17:27:57
Job time : 1826 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Om protein - protein search, using sw model

Run on: June 9, 2005, 17:28:08 ; Search time 94 Seconds
(without alignments)
3295.692 Million cell updates/sec

Title: US-10-600-070-2

Perfect score: 4063

Sequence: 1 MEALSHYIGISLPQLCRP.....YEVFWSKSGMKITRGVTLAS 801

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4063	100.0	801	ADJ38202	Adj38202 Arabidops
2	4051	99.7	801	ADJ38207	Adj38207 Placitid d
3	4051	99.7	801	ADJ38211	Adj38211 Placitid d
4	4051	99.7	801	ADJ38209	Adj38209 Placitid d
5	1775.5	43.7	760	ADJ38205	Adj38205 Placitid d
6	1646	40.5	324	ADJ38203	Adj38203 Arabidops
7	524.5	12.9	768	ADJ38247	Adj38247 Placitid d
8	498.5	12.3	789	ADJ38274	Adj38274 Placitid d
9	497	12.2	798	ADJ38244	Adj38244 Placitid d
10	497	12.2	798	ADJ38245	Adj38245 Placitid d
11	442	10.9	631	ADJ38242	Adj38242 Placitid d
12	442	10.9	631	ADJ38132	Adj38132 Synchoco
13	442	10.9	631	ADJ38240	Adj38240 Placitid d
14	431	10.6	819	ADJ38253	Adj38253 Placitid d
15	431	10.6	819	ADJ38251	Adj38251 Placitid d
16	429	10.6	714	ADJ38249	Adj38249 Placitid d
17	429	10.6	714	ADJ38250	Adj38250 Placitid d
18	423	10.4	673	ADJ38272	Adj38272 Placitid d
19	413	10.2	836	ADJ38270	Adj38270 Placitid d
20	401	9.9	716	ADJ38239	Adj38239 Placitid d
21	344.5	8.5	702	ADJ38235	Adj38235 Placitid d
22	320	7.9	661	ADJ38237	Adj38237 Placitid d
23	145.5	3.6	6095	AAU10701	Amino aci
24	143	3.5	2089	AAW08333	Cyclolell
25	139.5	3.4	954	AAW72752	DNA polym

26	139.5	3.4	954	2	ADJ87484	Adj87484 Confectio
27	138	3.3	2021	7	ADJ70511	Adj70511 Human hea
28	133.5	3.3	6973	7	ADJ26982	Adj26982 Sorangium
29	133	3.3	1273	4	ABG00113	Abg00113 Novel hum
30	132.5	3.3	876	8	ADM13791	Adm13791 Bacillus
31	131.5	3.2	2111	4	AAB66471	Aab66471 Protein e
32	130.5	3.2	876	6	ABP43335	Abp43335 Bacillus
33	130	3.2	1082	5	ABP62883	Abp62883 Human pol
34	128.5	3.2	876	2	AAR80137	Aar80137 A. steatoc
35	128.5	3.2	876	2	AAR80143	Aar80143 B. steatoc
36	128.5	3.2	876	2	AAR80142	Aar80142 B. steatoc
37	128	3.2	9234	7	ADC26980	Adc26980 Sorangium
38	127.5	3.1	2472	4	AAU33568	Aau33568 Pseudomon
39	127	3.1	1241	3	AAV93337	Aav93337 A human i
40	127	3.1	1241	6	ADA14856	Ada14856 Human int
41	127	3.1	1241	7	ABR42349	Abt42349 Human int
42	127	3.1	1241	8	ADR46554	Adt46554 Cancer-as
43	126.5	3.1	1169	5	ABP29745	Abp29745 Streptoco
44	126.5	3.1	1224	5	ABP26662	Abp26662 Streptoco
45	126.5	3.1	2482	7	ABO68221	Abp68221 Pseudomon

ALIGNMENTS

RESULT 1	
ADJ38202	ADJ38202 standard; protein; 801 AA.
XX	XX
AC	ADJ38202;
XX	XX
DT	06-MAY-2004 (first entry)
XX	XX
DE	Arabidopsis thaliana Arc6-1 protein.
XX	XX
KW	prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KW	agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW	herbicide target.
XX	XX
OS	Arabidopsis thaliana.
XX	XX
PN	WO2004001003-A2.
XX	XX
PD	31-DEC-2003.
XX	XX
FP	20-JUN-2003; 2003WO-US019536.
XX	XX
PR	20-JUN-2002; 2002US-0390140P.
XX	XX
PR	09-AUG-2002; 2002US-040242P.
XX	XX
PR	20-JUN-2003; 2003US-00600070.
XX	XX
PA	(UNMS) UNIV MICHIGAN STATE.
XX	XX
PI	Osteryoung KW, Vittha S, Kokscharova OA, Gao H;
XX	XX
DR	WPI: 2004-082486/08.
XX	XX
PT	N-PSDB; ADJ38129, ADJ38130.
XX	XX
PT	New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
XX	XX
PT	for further characterizing plastid division in plant cells, and in
XX	XX
PT	varying agronomic and horticultural characteristics of economically
XX	XX
PS	important plants.
XX	XX
PS	Disclosure; Fig 2, 287pp; English.
XX	XX
CC	This invention relates to novel prokaryotic type or plastid division and
XX	XX
CC	related genes and proteins. In particular, the invention relates to novel
XX	XX
CC	Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
XX	XX
CC	compositions of the present invention are useful for further
XX	XX
CC	characterizing plastid division in plant cells, in order to vary
XX	XX
CC	agronomic and horticultural characteristics of economically important
XX	XX
CC	plants, such as crop, ornamental and woody plants. They can also be used
XX	XX
CC	as herbicide targets. The present sequence is that of a protein which is

CC related to the invention.

XX Sequence 801 AA;

Query Match 100.0%; Score 4063; DB 8; Length 801;

Best Local Similarity 100.0%; Pred. No. 0; Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MEALSHVIGIGLSPFQOLCRLPATTKLRSHNTSTTICSAKMDRLSDPNFTSDSSSS 60
DB 1 MEALSHVIGIGLSPFQOLCRLPATTKLRSHNTSTTICSAKMDRLSDPNFTSDSSSS 60
QY 61 PATATTATTVSLPSPSIDRPERHVPIDFYQVLAQTHFLTGIRAFARVSKPPQFG 120
DB 61 PATATTATTVSLPSPSIDRPERHVPIDFYQVLAQTHFLTGIRAFARVSKPPQFG 120
QY 121 FSDDALISRQIIOACETLSNPRSRRYNEGILLDEBEATVITDVPMDKVPALCVLOEG 180
DB 121 FSDDALISRQIIOACETLSNPRSRRYNEGILLDEBEATVITDVPMDKVPALCVLOEG 180
QY 181 GETEIVLRVGEALLKERLPKSFQODVVLVMAALFLDVSRDAMALDPPDFTGYEFVEAL 240
DB 181 GETEIVLRVGEALLKERLPKSFQODVVLVMAALFLDVSRDAMALDPPDFTGYEFVEAL 240
QY 241 KLIQEGASSLAPDLRAQIDETLEITPRVYLLGLPLGDDVYAKRLNGLSGVRNITLMS 300
DB 241 KLIQEGASSLAPDLRAQIDETLEITPRVYLLGLPLGDDVYAKRLNGLSGVRNITLMS 300
QY 301 VGGGASALVGGLTREKFMNEAFLRMTAAEQVDLFVATPSNIPAESFEVEVALVAQA 360
DB 301 VGGGASALVGGLTREKFMNEAFLRMTAAEQVDLFVATPSNIPAESFEVEVALVAQA 360
QY 361 FIGKPHLLDADKQFOLOQAQKVMAMEIPALMYDTNNWEIDBLERGLCALLIGKVD 420
DB 361 FIGKPHLLDADKQFOLOQAQKVMAMEIPALMYDTNNWEIDBLERGLCALLIGKVD 420
QY 421 CRMWLGIDSEDSQYRNPAIVEFVLNSNRDNDLPGICLLETWLAGVFPFRPRTDK 480
DB 421 CRMWLGIDSEDSQYRNPAIVEFVLNSNRDNDLPGICLLETWLAGVFPFRPRTDK 480
QY 481 KFKLGDIYDDPMVLSYLERVEVVGSPLAATAATMARGAEHVKASAMQALQKVPSTTD 540
DB 481 KFKLGDIYDDPMVLSYLERVEVVGSPLAATAATMARGAEHVKASAMQALQKVPSTTD 540
QY 541 RNSAEPDVOETVESVDPVGNVNRDGPVFIAEAVRPSSENFETNDYAIRAGVSESSVD 600
DB 541 RNSAEPDVOETVESVDPVGNVNRDGPVFIAEAVRPSSENFETNDYAIRAGVSESSVD 600
QY 601 ETTVEMSVADMKEASVKILAAVAIGLISLFSQKTFPKSSSFQKQDVMSSMESDVATI 660
DB 601 ETTVEMSVADMKEASVKILAAVAIGLISLFSQKTFPKSSSFQKQDVMSSMESDVATI 660
QY 661 GSVRADSESLPRMDATTAENIVSKWKIKSLAGPGRILEMLDGMRLKWTBRAA 720
DB 661 GSVRADSESLPRMDATTAENIVSKWKIKSLAGPGRILEMLDGMRLKWTBRAA 720
QY 721 ETQOLGLVYDTLLKLKSVDSVTVSADGTALVEATLESACLSDLVHPENNAITVRYTTT 780
DB 721 ETQOLGLVYDTLLKLKSVDSVTVSADGTALVEATLESACLSDLVHPENNAITVRYTTT 780
QY 781 RYEVFMKSGWKITGGSVLAS 801
DB 781 RYEVFMKSGWKITGGSVLAS 801

```

RESULT 2

ID ADJ38207 standard; protein; 801 AA.

XX AC ADJ38207;

DT 06-MAY-2004 (first entry)

XX

DE Plastid division-related Arc6 protein orthologue 2.

KW prokaryotic type; plastid division; Ftn2; Arc6; Arc5; Fzo; plant cell;

KM agronomic; horticultural; crop plant; ornamental plant; woody plant;

XX herbicide target.

OS Arabidopsis thaliana.

XX WO2004001003-A2.

PD 31-DEC-2003.

XX 20-JUN-2003; 2003WO-US019536.

XX 20-JUN-2002; 2002US-0390140P.

XX 09-AUG-2002; 2002US-0402242P.

XX 20-JUN-2003; 2003US-00600070.

XX (UNMS) UNIV MICHIGAN STATE.

XX Oesteryoung KW, Vicha S, Koksharova OA, Gao H;

XX WPI; 2004-082486/08.

XX N-PSDB; ADJ38208.

XX New isolated Ftn2, Arc5 and/or Fzo-like nucleic acid sequences, useful

XX for further characterizing plastid division in plant cells, and in

XX varying agronomic and horticultural characteristics of economically

XX important plants.

XX Diaclosure; Fig 8; 287bp; English.

XX This invention relates to novel prokaryotic type or plastid division

XX related genes and proteins. In particular, the invention relates to novel

XX Ftn2 (Arc6), Arc5 and Fzo-like genes and polypeptides. The methods and

XX compositions of the present invention are useful for further

XX characterizing plastid division in plant cells, in order to vary

XX agronomic and horticultural characteristics of economically important

XX plants, such as crop, ornamental and woody plants. They can also be used

XX as herbicide targets. The present sequence is that of a protein which is

XX related to the invention.

XX Sequence 801 AA;

XX Query Match 99.7%; Score 4051; DB 8; Length 801;

XX Best Local Similarity 99.8%; Pred. No. 0;

XX Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEALSHVIGIGLSPFQOLCRLPATTKLRSHNTSTTICSAKMDRLSDPNFTSDSSSS 60

DB 1 MEALSHVIGIGLSPFQOLCRLPATTKLRSHNTSTTICSAKMDRLSDPNFTSDSSSS 60

QY 61 PATATTATTVSLPSPSIDRPERHVPIDFYQVLAQTHFLTGIRAFARVSKPPQFG 120

DB 61 PATATTATTVSLPSPSIDRPERHVPIDFYQVLAQTHFLTGIRAFARVSKPPQFG 120

QY 121 FSDDALISRQIIOACETLSNPRSRRYNEGILLDEBEATVITDVPMDKVPALCVLOEG 180

DB 121 FSDDALISRQIIOACETLSNPRSRRYNEGILLDEBEATVITDVPMDKVPALCVLOEG 180

QY 181 GETEIVLRVGEALLKERLPKSFQODVVLVMAALFLDVSRDAMALDPPDFTGYEFVEAL 240

DB 181 GETEIVLRVGEALLKERLPKSFQODVVLVMAALFLDVSRDAMALDPPDFTGYEFVEAL 240

QY 241 KLIQEGASSLAPDLRAQIDETLEITPRVYLLGLPLGDDVYAKRLNGLSGVRNITLMS 300

DB 241 KLIQEGASSLAPDLRAQIDETLEITPRVYLLGLPLGDDVYAKRLNGLSGVRNITLMS 300

QY 301 VGGGASALVGGLTREKFMNEAFLRMTAAEQVDLFVATPSNIPAESFEVEVALVAQA 360

DB 301 VGGGASALVGGLTREKFMNEAFLRMTAAEQVDLFVATPSNIPAESFEVEVALVAQA 360

QY 361 FIGKPHLLDADKQFOLOQAQKVMAMEIPALMYDTNNWEIDBLERGLCALLIGKVD 420

DB 361 FIGKPHLLDADKQFOLOQAQKVMAMEIPALMYDTNNWEIDBLERGLCALLIGKVD 420

```

Db      361 FIGKKPHLLQDADKQFQOLQOAKMAMEIPAMLYDTRNNWEIDPGERGICALLIGKVD 420
Qy      421 CRMWLGIDSDSQYRNPAIVEFPLENSNRDNDLPGLCLETLWLAGVFPFRDITDK 480
Db      421 CRMWLGIDSDSQYRNPAIVEFPLENSNRDNDLPGLCLETLWLAGVFPFRDITDK 480
Qy      481 KFKLGDYDDPMTLSYERVEVQGSPLAATAATMARIGAEHVKASAMQALQKVPFSRYTD 540
Db      481 KFKLGDYDDPMTLSYERVEVQGSPLAATAATMARIGAEHVKASAMQALQKVPFSRYTD 540
Qy      541 RNSAEPDQVEITYFSDVPVGNVNGRDGEPGVFIAEAVRPESENFETNDYAIRAGVSESSVD 600
Db      541 RNSAEPDQVEITYFSDVPVGNVNGRDGEPGVFIAEAVRPESENFETNDYAIRAGVSESSVD 600
Qy      601 ETTVMSVADMKEASVKIIAAGVAILGLISLFSQKFLKSSSFORKDMWSMESDVATI 660
Db      601 ETTVMSVADMKEASVKIIAAGVAILGLISLFSQKFLKSSSFORKDMWSMESDVATI 660
Qy      661 GSVRADDSSEALPRMDARTAEINIVSKMOKIKSLAFGPDHRIEMLPEVLDGRLKIWTDRAA 720
Db      661 GSVRADDSSEALPRMDARTAEINIVSKMOKIKSLAFGPDHRIEMLPEVLDGRLKIWTDRAA 720
Qy      721 ETTAQLGLVDDYTLKLKLSVDSVTVSADGTRALVEATLEESACLSGLVHPENNATDVRTYTT 780
Db      721 ETTAQLGLVDDYTLKLKLSVDSVTVSADGTRALVEATLEESACLSGLVHPENNATDVRTYTT 780
Qy      781 RYEVFWSKSGMKITTEGSVLAS 801
Db      781 RYEVFWSKSGMKITTEGSVLAS 801

```

RESULT 3

ID ADJ38211 standard; protein; 801 AA.

AC ADJ38211;

XX 06-MAY-2004 (first entry)

DE Placaid division-related Arc6 protein orthologue 4.

XX prokaryotic type; Placaid division; Ftn2; ARC6; ARCS; Fzo; plant cell;
 XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
 XX herbicide target.

XX Arabidopsis thaliana.

XX WO2004001003-A2.

XX 31-DEC-2003.

XX 20-JUN-2003; 2003WO-US019536.

XX 20-JUN-2002; 2002US-0390140P.

XX 09-AUG-2002; 2002US-0402242P.

XX 20-JUN-2003; 2003US-0060070.

XX (UNMS) UNIV MICHIGAN STATE.

XX Oesteryoung KM, Vitcha S, Koksharova OA, Gao H;

XX WPI; 2004-082486/08.

XX New isolated Ftn2, ARCS and Fzo-like nucleic acid sequences, useful
 XX for further characterizing placid division in plant cells, and in
 XX varying agronomic and horticultural characteristics of economically
 XX important plants.

XX Disclosure; Fig 8; 287bp; English.

XX This invention relates to novel prokaryotic type or placid division and
 XX related genes and proteins. In particular, the invention relates to novel

CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing placid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a protein which is
 CC related to the invention.

XX Sequence 801 AA;

Query Match 99.7%; Score 4051; DB 8; Length 801;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy      1 MEALSHVIGLSLSPOLCRPLPATTKLRRSHNTSTTICSAKMDRLSDNFTSDSSSS 60
Db      1 MEALSHVIGLSLSPOLCRPLPATTKLRRSHNTSTTICSAKMDRLSDNFTSDSSSS 60
Qy      61 FATATTTATLVSIPSIDRERHVPIDFYQVLGAQTHELTGIRRAFEARVSKPPQFG 120
Db      61 FATATTTATLVSIPSIDRERHVPIDFYQVLGAQTHELTGIRRAFEARVSKPPQFG 120
Qy      121 FSDDALISRQIIQAACETLSNPRSRREYEGLLDDEBATVITDVPMDKVPGLCVLQEG 180
Db      121 FSDDALISRQIIQAACETLSNPRSRREYEGLLDDEBATVITDVPMDKVPGLCVLQEG 180
Qy      181 GETEIVLRGEALTKERLPKSFQODVVLVWALFLDVSRRMALDPPDFITGYEFVEAL 240
Db      181 GETEIVLRGEALTKERLPKSFQODVVLVWALFLDVSRRMALDPPDFITGYEFVEAL 240
Qy      241 KLLQEBGASSLAPDLRAQIDETLEITPRVYLELGLPLGDDYAAKRLNGLSGRNITLWS 300
Db      241 KLLQEBGASSLAPDLRAQIDETLEITPRVYLELGLPLGDDYAAKRLNGLSGRNITLWS 300
Qy      301 VGGGASALVVGITREKFNNEATLRTAEOVDLFWATPSNIPAESEVEYVALVAQA 360
Db      301 VGGGASALVVGITREKFNNEATLRTAEOVDLFWATPSNIPAESEVEYVALVAQA 360
Qy      361 FIGKKPHLLQDADKQFQOLQOAKMAMEIPAMLYDTRNNWEIDPGERGICALLIGKVD 420
Db      361 FIGKKPHLLQDADKQFQOLQOAKMAMEIPAMLYDTRNNWEIDPGERGICALLIGKVD 420
Qy      421 CRMWLGIDSDSQYRNPAIVEFPLENSNRDNDLPGLCLETLWLAGVFPFRDITDK 480
Db      421 CRMWLGIDSDSQYRNPAIVEFPLENSNRDNDLPGLCLETLWLAGVFPFRDITDK 480
Qy      481 KFKLGDYDDPMTLSYERVEVQGSPLAATAATMARIGAEHVKASAMQALQKVPFSRYTD 540
Db      481 KFKLGDYDDPMTLSYERVEVQGSPLAATAATMARIGAEHVKASAMQALQKVPFSRYTD 540
Qy      541 RNSAEPDQVEITYFSDVPVGNVNGRDGEPGVFIAEAVRPESENFETNDYAIRAGVSESSVD 600
Db      541 RNSAEPDQVEITYFSDVPVGNVNGRDGEPGVFIAEAVRPESENFETNDYAIRAGVSESSVD 600
Qy      601 ETTVMSVADMKEASVKIIAAGVAILGLISLFSQKFLKSSSFORKDMWSMESDVATI 660
Db      601 ETTVMSVADMKEASVKIIAAGVAILGLISLFSQKFLKSSSFORKDMWSMESDVATI 660
Qy      661 GSVRADDSSEALPRMDARTAEINIVSKMOKIKSLAFGPDHRIEMLPEVLDGRLKIWTDRAA 720
Db      661 GSVRADDSSEALPRMDARTAEINIVSKMOKIKSLAFGPDHRIEMLPEVLDGRLKIWTDRAA 720
Qy      721 ETTAQLGLVDDYTLKLKLSVDSVTVSADGTRALVEATLEESACLSGLVHPENNATDVRTYTT 780
Db      721 ETTAQLGLVDDYTLKLKLSVDSVTVSADGTRALVEATLEESACLSGLVHPENNATDVRTYTT 780
Qy      781 RYEVFWSKSGMKITTEGSVLAS 801
Db      781 RYEVFWSKSGMKITTEGSVLAS 801

```

RESULT 4

ADJ38209

```

ID ADJ38209 standard; protein; 801 AA.
XX
XX ADJ38209;
XX
XX 06-MAY-2004 (first entry)
XX
XX plastid division-related Arc6 protein orthologue 3.
XX
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target.
XX
XX Arabidopsis thaliana.
XX
XX W02004001003-A2.
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003MO-US019536.
XX
XX 20-JUN-2002; 2002US-0390140P.
XX PR 09-AUG-2002; 2002US-040242P.
XX PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS ) UNIV MICHIGAN STATE.
XX
XX Oosteryoung KW, Viltha S, Kokeharova OA, Gao H;
XX
XX WPI; 2004-082486/08.
XX
XX N-PSDB; ADJ38210.
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX PT for further characterizing plastid division in plant cells, and in
XX PT varying agronomic and horticultural characteristics of economically
XX PT important plants.
XX
XX Disclosure; Fig 8; 287bp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
XX CC related genes and proteins. In particular, the invention relates to novel
XX CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
XX CC compositions of the present invention are useful for further
XX CC characterizing plastid division in plant cells, in order to vary
XX CC agronomic and horticultural characteristics of economically important
XX CC plants, such as crop, ornamental and woody plants. They can also be used
XX CC as herbicide targets. The present sequence is that of a protein which is
XX CC related to the invention.
XX
XX Sequence 801 AA;
XX
XX Query Match 99.7%; Score 4051; DB 8; Length 801;
XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 MEALSHVIGIGLSPFQLCRLPATTKLRSHNTSTTTCSGAKMADRLSDNFNTSSSSSS 60
XX 1 MEALSHVIGIGLSPFQLCRLPATTKLRSHNTSTTTCSGAKMADRLSDNFNTSSSSSS 60
XX
XX 61 FATATTATVLSPESTIDRBRHVPIPIDFYOVLAGQTHFLTGGIRAFEARVSKPQFG 120
XX 61 FATATTATVLSPESTIDRBRHVPIPIDFYOVLAGQTHFLTGGIRAFEARVSKPQFG 120
XX
XX 61 FATATTATVLSPESTIDRBRHVPIPIDFYOVLAGQTHFLTGGIRAFEARVSKPQFG 120
XX 61 FATATTATVLSPESTIDRBRHVPIPIDFYOVLAGQTHFLTGGIRAFEARVSKPQFG 120
XX
XX 121 FSDDALISRRQIIQACETLSNPRSREYNEGGLDDEEATVITDVPMVKYFALCVLQEG 180
XX 121 FSDDALISRRQIIQACETLSNPRSREYNEGGLDDEEATVITDVPMVKYFALCVLQEG 180
XX
XX 121 FSDDALISRRQIIQACETLSNPRSREYNEGGLDDEEATVITDVPMVKYFALCVLQEG 180
XX 121 FSDDALISRRQIIQACETLSNPRSREYNEGGLDDEEATVITDVPMVKYFALCVLQEG 180
XX
XX 181 GETEIVIRVEBALIKERLPKSPKQDVVLVMAALFADVSBDAMALDPDFITGYEFVEAL 240
XX 181 GETEIVIRVEBALIKERLPKSPKQDVVLVMAALFADVSBDAMALDPDFITGYEFVEAL 240
XX
XX 181 GETEIVIRVEBALIKERLPKSPKQDVVLVMAALFADVSBDAMALDPDFITGYEFVEAL 240
XX 181 GETEIVIRVEBALIKERLPKSPKQDVVLVMAALFADVSBDAMALDPDFITGYEFVEAL 240
XX
XX 241 KLOEBGASLAPDLRAQIDETLEETPRVYELILGPIGDVYAAKRLNGISGRNLTMS 300
XX 241 KLOEBGASLAPDLRAQIDETLEETPRVYELILGPIGDVYAAKRLNGISGRNLTMS 300
XX
XX 241 KLOEBGASLAPDLRAQIDETLEETPRVYELILGPIGDVYAAKRLNGISGRNLTMS 300
XX 241 KLOEBGASLAPDLRAQIDETLEETPRVYELILGPIGDVYAAKRLNGISGRNLTMS 300

```

```

QY 301 VGGGASALVGLTREKFNNEAFLRMTAAEQVDLFAVTPSNIPIAESFEVEVALALVAQA 360
QY 301 VGGGASALVGLTREKFNNEAFLRMTAAEQVDLFAVTPSNIPIAESFEVEVALALVAQA 360
DB 301 VGGGASALVGLTREKFNNEAFLRMTAAEQVDLFAVTPSNIPIAESFEVEVALALVAQA 360
QY 361 FIKKKPHLLQDADKQOQOQAAKVMAMEIPAMLYDRNMNEIDFGIERGICALLIGKVD 420
QY 361 FIKKKPHLLQDADKQOQOQAAKVMAMEIPAMLYDRNMNEIDFGIERGICALLIGKVD 420
DB 361 FIKKKPHLLQDADKQOQOQAAKVMAMEIPAMLYDRNMNEIDFGIERGICALLIGKVD 420
QY 421 CRMWLGIDSDSQYRNPAIVEFLIENSNRDNDLPGLCLEETWLAGVFPFRFRTDKX 480
QY 421 CRMWLGIDSDSQYRNPAIVEFLIENSNRDNDLPGLCLEETWLAGVFPFRFRTDKX 480
DB 421 CRMWLGIDSDSQYRNPAIVEFLIENSNRDNDLPGLCLEETWLAGVFPFRFRTDKX 480
QY 481 KFKLGDYDDPMVLSTYLERVVOGSPPLAAATMARI GAHNVKASAMQALQKVPSTRTD 540
QY 481 KFKLGDYDDPMVLSTYLERVVOGSPPLAAATMARI GAHNVKASAMQALQKVPSTRTD 540
DB 481 KFKLGDYDDPMVLSTYLERVVOGSPPLAAATMARI GAHNVKASAMQALQKVPSTRTD 540
QY 541 RNSAEPDQVETVPSVDPVGNNGRDEPGVFIAEAVRPSNEFTNDYAIRAVSSSSVD 600
QY 541 RNSAEPDQVETVPSVDPVGNNGRDEPGVFIAEAVRPSNEFTNDYAIRAVSSSSVD 600
DB 541 RNSAEPDQVETVPSVDPVGNNGRDEPGVFIAEAVRPSNEFTNDYAIRAVSSSSVD 600
QY 601 ETTVMSVADMLKEASVKILAAAGVAILGLISLFSQKFLKSSSFORKDMVSSMESDVATI 660
QY 601 ETTVMSVADMLKEASVKILAAAGVAILGLISLFSQKFLKSSSFORKDMVSSMESDVATI 660
DB 601 ETTVMSVADMLKEASVKILAAAGVAILGLISLFSQKFLKSSSFORKDMVSSMESDVATI 660
QY 661 GSVRADSEALPMDARTAEINIVSKWQIKSLAFGPDHRIEMLPVLDGRLKIWTDRAA 720
QY 661 GSVRADSEALPMDARTAEINIVSKWQIKSLAFGPDHRIEMLPVLDGRLKIWTDRAA 720
DB 661 GSVRADSEALPMDARTAEINIVSKWQIKSLAFGPDHRIEMLPVLDGRLKIWTDRAA 720
QY 721 ETRQDLGVYDYTLTKLSVSVSVYASADGTRALVARTLEESACSLDVHPENNATVRYTTT 780
QY 721 ETRQDLGVYDYTLTKLSVSVSVYASADGTRALVARTLEESACSLDVHPENNATVRYTTT 780
DB 721 ETRQDLGVYDYTLTKLSVSVSVYASADGTRALVARTLEESACSLDVHPENNATVRYTTT 780
QY 781 RYEVFMSKSGMKITTEGSLAS 801
QY 781 RYEVFMSKSGMKITTEGSLAS 801
DB 781 RYEVFMSKSGMKITTEGSLAS 801

```

```

RESULT 5
ADJ38205
ID ADJ38205 standard; protein; 760 AA.
XX
XX ADJ38205;
XX
XX 06-MAY-2004 (first entry)
XX
XX plastid division-related Arc6 protein orthologue 1.
XX
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
XX KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX KW herbicide target.
XX
XX Oryza sativa.
XX
XX W02004001003-A2.
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003MO-US019536.
XX
XX 20-JUN-2002; 2002US-0390140P.
XX PR 09-AUG-2002; 2002US-040242P.
XX PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS ) UNIV MICHIGAN STATE.
XX
XX Oosteryoung KW, Viltha S, Kokeharova OA, Gao H;
XX
XX WPI; 2004-082486/08.
XX
XX N-PSDB; ADJ38206.
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX PT for further characterizing plastid division in plant cells, and in

```

PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287pp; English.
 CC This invention relates to novel prokaryotic type or plasmid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plasmid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a protein which is
 CC related to the invention.
 XX
 SQ Sequence 760 AA;
 Query Match 43.7%; Score 1775.5; DB 8; Length 760;
 Best Local Similarity 47.9%; Pred. No. 1.9e-150;
 Matches 390; Conservative 119; Mismatches 213; Indels 93; Gaps 16;
 QY 12 SPFQCLRPPTTKLRSHNTSTTC-SASKMDRLSDNF-----TSQSSSSFRATAT 65
 DB 14 APFAPSPPRRPRRRPPPPPSACRAASRAERLPADFHLPTAAPSDPPSPAPAPAA 73
 QY 66 TTATLVSLPSPIDRERHVPPIPIFYQVLAQTHFLTDGIRAFARVSKPPQFGSDA 125
 DB 74 APSASPPVPLPPDAERSLPLOVPYKVLGAPHLGDGIRAFARVSKPPQFGSTDA 133
 QY 126 LISRRQILQAACETLSNPSRREYNEGLDDEEATVITDVPMDKVPGLCVLOEGGETEI 185
 DB 134 LVGRKQLQIAHDITLNMQNSRTQYDRALSENREELMTDMDK-----EAGEMLA 184
 QY 186 VLRVGAALLERLPEKSKQDVLYMALAFDVSSDAMALDPPDTITGYEYERALKLQI 245
 DB 185 VLVYGEQLLIDRPKRFKQDVLYMALAYDLISRDAMASPPDVIGCEVEYERALKLQI 244
 QY 246 EGASSLAPDLRAOIDEETLEETTPRYVELGLPIGDGDAARLMLGSGVRITLVSGGG 305
 DB 245 DGSNMLPDLISQIDEETLEETTPRCVLESLPDTBHHKROGLQANRILMVSGRG 304
 QY 306 ASALVGLTEKEFNNEAFLEMTAAEQVDLFVATPSNIPAESFEYEVALLVAQAFLGKK 365
 DB 305 IATVGGFSREAFNEAFLEMTAAEQVDLFVATPSNIPAESFEYEVALLVAQAFLGKK 364
 QY 366 PHLIQDDKPOQLQAQKWMAMEIPAMLYTRNNWEIDFGLERGLCALLIGKVDCEKWL 425
 DB 365 PQFIMMADDLFEQLQKFNIGS---HYAYDN---EMDLALERAFCSLVGVDSKCRMWL 416
 QY 426 GLDSEDSQYRNPAIVEFVLENSN-RDNDNDLPGICXLEETLAVGVPRPDTDKKPKL 484
 DB 417 GIDNESSPQKLEFVLTNSSLSEENDLPGICXLEETLAVGVPRPDTDKKPKL 476
 QY 485 GDYDDDDPVLSTYLERVEYQGSPLAAATMARIGAE-----HYKASMQALQKVP-SP 537
 DB 477 GDYDDDDPVLSTYLERVEYQGSPLAAATMARIGAE-----HYKASMQALQKVP-SP 536
 QY 538 YTDNRSAEPKQDVQTFVSDVGVNNGRGGEPGVFLAEAVRPSNFTNDYAIRAGVSES 597
 DB 537 QLDPSAMENT-----KDG-PGGYL-----BNFQO----- 559
 QY 598 SVBETTEMASVADMLKASVYKILAGVALIGLSFSQKYPFKSSSSQKQKMSMSD- 656
 DB 560 -----ENAFHDSRNPAALKISAGALFALAVYGAET-----LPRKGPLARISEH 605
 QY 657 --VATIGSVADDBEAL-----PRMDARTENIVSKQIKSLAFGPDHRIEMLEPVL 707
 DB 606 GSVAVASVSTDDPALDEDPVHLPRMDAKLABDIVRMWSIKSKALGEPHSVASLQEVL 665
 QY 708 DGRMLKWTDBRAETAQGLVYDYTLKLSDVSTVVSADGTRALVEATLEESACLSDLVH 767
 DB 666 DGNMLKWTDBRAETIERHGFMEYTLSDVITDSITISLIDGRRAVTEATIDEGQLTDTVE 725

QY 768 PENNATDVRTYTRYEVFWSK-SGMKITEGSVLAS 801
 DB 726 PRNNDSDYTKTYTRYEMAFSKLGSMTITBGAVLKS 760
 RESULT 6
 ADJ38203
 ID ADJ38203 standard; protein; 324 AA.
 XX
 AC ADJ38203;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Arabidopsis thaliana AtFtn2 mutated protein.
 XX
 KM prokaryotic type; plasmid division; Ftn2; ARCS; ARCS; Fzo; plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KM herbicide target; mutant; mutain.
 XX
 OS Arabidopsis thaliana.
 OS
 PN W02004001003-A2.
 PN
 PD 31-DEC-2003.
 XX
 PF 20-JUN-2003; 2003W0-US019536.
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Oosteryoung KM, Vitsha S, Kosharova OA, Gao H;
 XX
 DR WPI; 2004-082486/08.
 DR N-PSDB; ADJ38135, ADJ38136.
 XX
 PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plasmid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 2; 287pp; English.
 CC This invention relates to novel prokaryotic type or plasmid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plasmid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a protein which is
 CC related to the invention.
 XX
 SQ Sequence 324 AA;
 Query Match 40.5%; Score 1646; DB 8; Length 324;
 Best Local Similarity 100.0%; Pred. No. 2.2e-139;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MEALSHVIGISLPOLCRLPATTKLRSHNTSTTCSASKMDRLSDNFSSSSSS 60
 DB 1 MEALSHVIGISLPOLCRLPATTKLRSHNTSTTCSASKMDRLSDNFSSSSSS 60
 QY 61 FATATTATVLSLPPSIDRERHVPPIPIFYQVLAQTHFLTDGIRAFARVSKPPQFG 120
 DB 61 FATATTATVLSLPPSIDRERHVPPIPIFYQVLAQTHFLTDGIRAFARVSKPPQFG 120
 QY 121 FSDDALISRRQILQAACETLSNPSRREYNEGLDDEEATVITDVPMDKVPGLCVLOEG 180
 DB 121 FSDDALISRRQILQAACETLSNPSRREYNEGLDDEEATVITDVPMDKVPGLCVLOEG 180

QY 181 GETEIVLRVGBALLKERLPKSFKODVVLVMAALFLVSRDMALDPDFITGYEFVEBAL 240
DB 181 GETEIVLRVGBALLKERLPKSFKODVVLVMAALFLVSRDMALDPDFITGYEFVEBAL 240
QY 241 KLOEBSASSIAPDLRAIOETLEETPRVYLLELGLPLGDDVAAKXUNGISGVRNLTMS 300
DB 241 KLOEBSASSIAPDLRAIOETLEETPRVYLLELGLPLGDDVAAKXUNGISGVRNLTMS 300
QY 301 VGGGASALVGGLTREKPMNEAFL 324
DB 301 VGGGASALVGGLTREKPMNEAFL 324

RESULT 7
ADJ38247
ID ADJ38247 standard; protein; 768 AA.

AC ADJ38247;
XX 06-MAY-2004 (first entry)

DE Plastid division-related Arc6 orthologue protein 35.

XX prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target.

OS Unidentified.

XX W02004001003-A2.

XX 31-DEC-2003.

XX 20-JUN-2003; 2003WO-US019536.

XX 20-JUN-2002; 2002US-0390140P.

XX 09-AUG-2002; 2002US-040242P.

XX 20-JUN-2003; 2003US-0060070.

XX (UNMS) UNIV MICHIGAN STATE.

XX Oeteryoung KM, Vicha S, Kokeharova OA, Gao H;

XX WPI; 2004-082486/08.

XX N-PSDB; ADJ38246.

XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful

XX for further characterizing plastid division in plant cells, and in

XX varying agronomic and horticultural characteristics of economically

XX important plants.

XX Disclosure; Fig 8; 287pp; English.

XX This invention relates to novel prokaryotic type or plastid division and

XX related genes and proteins. In particular, the invention relates to novel

XX Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and

XX compositions of the present invention are useful for further

XX characterizing plastid division in plant cells, in order to vary

XX agronomic and horticultural characteristics of economically important

XX plants, such as crop, ornamental and woody plants. They can also be used

XX as herbicide targets. The present sequence is that of a protein which is

XX related to the invention.

XX Sequence 768 AA;

XX Query Match 12.9%; Score 524.5; DB 8; Length 768;

XX Best Local Similarity 24.9%; Pred. No. 1.66-37;

XX Matches 212; Conservative 128; Mismatches 286; Indels 227; Gaps 28;

XX 84 VPIPIDEYOVYGAQTHFLTDGIRAFARVSKPPOFGSDALISRRQILQACETLSNP 143

XX 1 VRIPUDYRILGLPLAASSEQLRQAYSDRIYQLRRRRTSQMAISSRKQLEAFVVLSDP 60

QY 144 RSRREINELGL-----DDEEATVIT-----DVPMDKPGALCVLQ 178
DB 61 KOSTYDQQLAYAYPDNIAAAYQENNTSTKSGSDTQSLGIEITDDELVALILQ 120
QY 179 EGGETEIVLRVGBALL--KERLPKSFQ-----DVVLVMAALFLVSRD- 220
DB 121 ELGSEYELVLRVGBALL--KERLPKSFQ-----DVVLVMAALFLVSRD- 220
QY 221 -----AMALDPDFITGYEFVEBALKLOEBSASSIAPDLRAIOETLEETPRVY 271
DB 181 MOGHYENMAISLE-----TGQ-----LVREG-----LFSSIQAEIQADLYKLRERYI 225

QY 272 LELLGLPLGDDVAAKXUNGISGVRNLTMSVG--CGASALVGGITRE--KPMNEAFLRMT 327
DB 226 LELLGLPLGDDVAAKXUNGISGVRNLTMSVG--CGASALVGGITRE--KPMNEAFLRMT 327

QY 328 AAEQVDFVATTPSNIPASFEVYEVVALVAQAFIGKPHLLDDADQFOQLQAAKMMAM 367
DB 284 VAEQHKLFEAQSRK--SSAVATYLAAYALTARQPAQAPALIRQAKOMLVRLGRQ--- 337

QY 388 EIPAMLYDTNNWEIDFGLEGLCALLIGKVCDECRMWLGDSQYRNPAIVEFVLENS 447
DB 338 -----DVHLEQSLCALLIGQTEATRLLELSQ---YE---ALAFIREKS 376

QY 448 NRDDNDLPELCLETWLAGVPPRRPRTYDKKFKLGDYDDPMVLSYERVAVQGGSP 507
DB 377 -QDSPLLPGLCLYAEQWLQHEVFPFHRDLANOQAFKDFANQOVAYALE----- 426

QY 508 LAAATATWARIQAEHVYASAMQALQVPPRYTNRNSAEPR-----DVOETVFSV 556
DB 427 --ALPTDAQTTNEMAYVNPQYFPOAKAKNTHFNNSKTTSASRHSKRVPPDLPEP- 480

QY 557 DPGNNVNGRDEBGFVIAEAVRPSSENEFTNDYAIRAGVSSSVDETTVENSVAQMLK-- 613
DB 481 -----PTK--ETSEYENFSPMWSGSIKSEVPALBRMSRGT 516

QY 614 ----EASVKTLAAG----- 623
DB 517 NOHLNGSAKSAASGHQKRRRRKPTSPASRERIIPDRPHSRPRRRRTFANTIEGKTRLV 576

QY 624 --VAIGLISL-----PSQKFLKSSSFQKMDVWSMESDVATIGSVRADSE 669
DB 577 WRVFLISVLSILVWVWVLAATTTFGWLNLFPPQBPPLQLFVQINQPPILPDPDRKRESE 636

QY 670 ALPRMDARTENIVSKQKIKSLAFGPDRIEMLPVLDGRMLKIWTDRAETPAQGLVY 729
DB 637 EGPLTNAE-AEEYIHWLSTKKAALGPNHINNLEQILTSALSQWR-LIAQNKLDNRY 694

QY 730 ---DYTLKLSDSVSVASDGTALYBATLEBACUSDLVHPENNATDVATTTTRVVFV 766
DB 695 RKFPHSLIKIESVERKIGLFAD--BAAVEATVKEVTOLEYENNOFQNSND--KLRRYDILR 750

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

QY 787 SKSGWKTEGGSVL 799
DB 751 ERGKMRIGSTSV 763

PN WO2004001003-A2.
 XX
 PD 31-DEC-2003.
 XX
 PE 20-JUN-2003; 2003WO-US019536.
 XX
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Oeseryoung KW, Vitha S, Koksharova OA, Gao H;
 XX
 DR WPI: 2004-082486/08.
 DR N-PSDB, ADJ38273.
 XX
 PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287pp; English.
 CC
 CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a protein which is
 CC related to the invention.
 CC
 SQ Sequence 789 AA;

Query Match 12.3%; Score 498.5; DB 8; Length 789;
 Best Local Similarity 23.7%; Pred. No. 3.8e-35;
 Matches 203; Conservative 132; Mismatches 306; Indels 215; Gaps 31;

QY 84 VPIPIIDYQVLAQTHLTDGIRAFARVSKPPQFGFSDALISRQILQACETLSNP 143
 DB 1 VRIPLDYRIRIGLPIQATBQRLQAHQDRTOQPPRRESEATYVARQGLIDEAVAVLCDF 60
 QY 144 RSREYNEGLI----- 154
 DB 61 EQRQTYDGNFLAKTYEPIVEELNPSQINPQAEKETTLEKREVLPEIASKOLKKRTS 120
 QY 155 -----DDEATVITDVPMKVPKALCVLQEGGETEIVLVGEA-LIKERL----- 198
 DB 121 YQNETYASDPHSNTPSIEIEYPOFGATILILHLEGYELVLTHTPYLNNSTITKDG 180
 QY 199 -----PKSFQDVVLVMAFLDVSMDAMALDPDPFITG-YEFVEALK-----LLQERGA 248
 DB 181 RFGDPALVLPVVLTVALANLELGR-----EMQGOYSATFALGAGLLIRE---- 230
 QY 249 SSIAADLRQAQIDETLEETPRYVLELLGLPGDDYAAARLNGLSGVNIIIVSG3-GGAS 307
 DB 231 -NLFPVQIRGEIQADLYKLRPYRIMELALP--BEIALDRSGRLTQMLNBERGIDGOG 287
 QY 308 ALVAGGLIRE---KMNNEAFLEMTAAEOVDLEVA---TPSNTPAESFEYVYVALVVAQAF 361
 DB 288 EDSGGIGLIDFLKTVQOQRLYTTABEQKLEAEALRSVAGV-----YLAVTFFLQGF 342
 QY 362 IGKPEHLLQDADKQFOLOQAKVWAMEIRPMLYDTRNMEIDFGLEGLCALLIGKYDEC 421
 DB 343 AQKQPAPIRKAKMLMQLGRSG-----DVNLKSVCLALLGQTEBA 383
 QY 422 RMLGLDSEDSQYRNPAIVFVLVENSNDNDLDPGLCKLFTWLAGVVPFRPDYDKK 481
 DB 384 SRSLSEHSENP-----LSFIKENSQOSP-DLPGCLTVAEHMLTBEVFPHPFDLSKDS 436
 QY 482 FKLDYDDPDPVLSYLERVEVVGSSPLAATAATMARIGAEHAKASAMQALQVFPFSRTYDR 541

DB 437 ASLDYFADQGVQAYLE-----ETVFSVDPVGNV---GRGEPGVFIAEAVRPEENFETDYA 478
 QY 542 NSAPKDVQ-----ETVFSVDPVGNV---GRGEPGVFIAEAVRPEENFETDYA 589
 DB 479 QMFDPKELKLVNSDLEDKDISRYDATATGIVASGQSSNLLGASSDGLQLEKSS-S 537
 QY 590 IRAG-----VSESSV-----DETVEM-SVADMLKEASVYK----- 618
 DB 538 TRGSPKQVTTKSSHYLGKIREKISGLPERNESTIESGGLPSQISGHSRRTSARRE 597
 QY 619 -----ILAGVA---IGLISLFSQKY--FLKSSSSFORKDMVSSMESDVATIGSVRAD 667
 DB 598 VKFGRLLILIAVGLTIGLITKITIGMLVNLGWEREKMTQLDRPPIEIPEDPRVN 657
 QY 668 SEALPRMDARTAEIVSKWKIKSLAGPDRHIMLEPVLDGRMLKTV--TDRAAEATQL 725
 DB 658 LAAGPITKEVARRTTISWDIKASALGPNNKIEQLNITLVEPALSRMLPYANLKOES 717
 QY 726 GLVYDYTL-----LKLSDVSVTVSADGTRALVEA-TLESASCTSD---LVHPENNATDVRT 777
 DB 718 YRREYHDLKISNITMS-----NTNSMLAQVDAAVIEKVEYSUNGRLTNTNENLPVRY 771
 QY 778 YTRREYEVFWSKSGWKI 793
 DB 772 DLVVKSGKMQISNMKV 787

RESULT 9
 ADJ38244
 ID ADJ38244 standard; protein; 798 AA.
 XX
 AC ADJ38244;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Plastid division-related Arc6 orthologue protein 33.
 XX
 KM prokaryotic type; plastid division; Ftn2, ARCS, ARCS, Fzo, plant cell;
 KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
 KM herbicide target.
 XX
 OS Anabaena sp.
 XX
 PN WO2004001003-A2.
 XX
 PD 31-DEC-2003.
 XX
 PE 20-JUN-2003; 2003WO-US019536.
 XX
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Oeseryoung KW, Vitha S, Koksharova OA, Gao H;
 XX
 DR WPI: 2004-082486/08.
 XX
 PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287pp; English.
 CC
 CC This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important

CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a protein which is
CC related to the invention.
XX

XX Sequence 798 AA:

Query Match 12.2%; Score 497; DB 8; Length 798;

Best Local Similarity 24.6%; Pred. No. 5.3e-35;
Matches 207; Conservative 139; Mismatches 301; Indels 196; Gaps 33;

```

QY 84 VPIPIDFYOVUGAOTFLTDGIRAFARVSKPPQGFSDALISROIIQAACETLSNP 143
DB 11 VRIPLDYRILGIPLAASDQLRQAYSRIVOLPRREYSQALASRKQLEAVVLSDP 70
QY 144 RSRREY-----NEGILDEEATVITDVPMDKVPGLCV 176
DB 71 KERSSYQOLYLAAHYDPDNATTKVAENRGDSNNGHFDVQSLSI--EVSSEELIGLLI 128
QY 177 LQEGETEIVLRVGEALL-----KERLPKSF-----KQDVVLVMAIAFLDVS 219
DB 129 LQELGEYELVTKIGRNYLGNQNGTASTRNGNHRTPPEFLDSSRPDILLTVALLASLELGR 188
QY 220 D-----AMALDPDPFITGFEVEEALKLQEBGASSLAPDLAQLDTELEITPR 269
DB 189 EQMOQGHYENAAALSLR-----TGQE-----VLFSEG--IPPSYQAEIQADLYKLRPY 233
QY 270 YVLELGLPIGDDYAAKRLNGLSGVNIIIMSVGG--GASALVGLTRE--KFMNEAFLR 325
DB 234 RIEELALP--QEKTIERHQGLDLSIDDRGIDGTGNDGSLNIDDLRFIFQQLRH 291
QY 326 MTAAEQVLDLFVATPPSNI PASFEEYVALALVAQAFIGKKPHLLQDADKQFOQLQAKVM 385
DB 292 LTVAEQHKLPFG--BSKRPS--AVATYLAIVYAS IARGFORQALIRHAKQILMRLSKQ-- 347
QY 386 AMELPAMLVTRNNWEIDFGIERGLCALLIGKVDCEMMGLDSEBQYRNPALVEEVL 445
DB 348 -----DVHLEOSLCALLIGQTEEATRVLELSQE--YENALAI-----R 383
QY 446 NSNRDNDMDLPGCLKLETWLAGVPRPRDTXOKKFKLGDYDDPMWLSYLERV----- 500
DB 384 EKSQDSFDLLPGLCLYAEQWLQNEVPHFRLSRQOASLKDYIFANQOVQVLEMLPDAE 443
QY 501 -----EVVQSSPLAAATMAR--IGAEN--VYASAMQALQKVPSS---RYTD 540
DB 444 TTMEWAVINRQSPQPRGNVSQSTPAKAPVGRANRGEASTRPVQRSHPSVNRQFH 503
QY 541 RNSAEPKVOGETYFSDVPVGNNGRDS-----PGVFIABAVRPSNFTND 587
DB 504 QNRTPPELPELTPETSNHRRPSSNFTTARENISTTDAYTDNYPPELIPVERASRP----- 555
QY 588 YAIRAGVS---ESSVDETV-----EMSVADMKEASVKILAA 622
DB 556 --VQPGVSGTQSTPPRQTPRRRRKKRQAVVNRGSHIHQORQSPSTLGRKTRILMIVL 613
QY 623 GVALIGLSLP---SQRV-FLKS-----SSSFQRKDWVSSMESDVATIGSVRADSEALPR 673
DB 614 G-SLGGILFMLVISTTFEGLKQNVFPAPPSLQGEQLSIQISQPELEIPDKNAQIQSPREVS 672
QY 674 MDARTANIVSKQOKISLAFGPDRHTEMPEVLDGMMKLTWDRAETQOLGLVITYTL 733
DB 673 LTEETARKIENMLATVASALGAEHKIESINEILITGSLQWMLIALQADNRHRRPS- 731
QY 734 LKISVSVTVS--ADGRALVEATLESAGLSDLVH--PENNAATDVRTYTRRYEAFVMSKG 790
DB 732 HSYKVDISIKSDIDPNPASVGAIVR--LTQFYENGQKGSSEDER--LVRVTELIRODDI 787
QY 791 WKI 793
DB 788 WRI 790

```

RESULT 10
ADJ38245

ID ADJ38245 standard; protein; 798 AA.

XX ADJ38245;

XX 06-MAY-2004 (first entry)

DE Plastid division-related Arc6 orthologue protein 34.

XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target.

XX Anabaena sp.

XX WO2004001003-A2.

XX 31-DEC-2003.

XX 20-JUN-2003; 2003WO-US019536.

XX 20-JUN-2002; 2002US-0390140P.

XX 09-AUG-2002; 2002US-0402242P.

XX 20-JUN-2003; 2003US-00600070.

XX (UNMS) UNIV MICHIGAN STATE.

XX Oosteryoung KW, Vitha S, Koksharova OA, Gao H;

XX WPI, 2004-08246/08.

XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful

XX for further characterizing plastid division in plant cells, and in

XX varying agronomic and horticultural characteristics of economically

XX important plants.

XX Disclosure, Fig 8; 287bp, English.

XX This invention relates to novel prokaryotic type or plastid division and

XX related genes and proteins. In particular, the invention relates to novel

XX Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and

XX compositions of the present invention are useful for further

XX characterizing plastid division in plant cells, in order to vary

XX agronomic and horticultural characteristics of economically important

XX plants, such as crop, ornamental and woody plants. They can also be used

XX as herbicide targets. The present sequence is that of a protein which is

XX related to the invention.

XX Sequence 798 AA:

XX Query Match 12.2%; Score 497; DB 8; Length 798;

XX Best Local Similarity 24.6%; Pred. No. 5.3e-35;

XX Matches 207; Conservative 139; Mismatches 301; Indels 196; Gaps 33;

QY 84 VPIPIDFYOVUGAOTFLTDGIRAFARVSKPPQGFSDALISROIIQAACETLSNP 143

DB 11 VRIPLDYRILGIPLAASDQLRQAYSRIVOLPRREYSQALASRKQLEAVVLSDP 70

QY 144 RSRREY-----NEGILDEEATVITDVPMDKVPGLCV 176

DB 71 KERSSYQOLYLAAHYDPDNATTKVAENRGDSNNGHFDVQSLSI--EVSSEELIGLLI 128

QY 177 LQEGETEIVLRVGEALL-----KERLPKSF-----KQDVVLVMAIAFLDVS 219

DB 129 LQELGEYELVTKIGRNYLGNQNGTASTRNGNHRTPPEFLDSSRPDILLTVALLASLELGR 188

QY 220 D-----AMALDPDPFITGFEVEEALKLQEBGASSLAPDLAQLDTELEITPR 269

DB 189 EQMOQGHYENAAALSLR-----TGQE-----VLFSEG--IPPSYQAEIQADLYKLRPY 233

QY 270 YVLELGLPIGDDYAAKRLNGLSGVNIIIMSVGG--GASALVGLTRE--KFMNEAFLR 325

DB 234 RIEELALP--QEKTIERHQGLDLSIDDRGIDGTGNDGSLNIDDLRFIFQQLRH 291

Db 582 AVQAQKVD-----VEQYVRGDQLLETRRDGLVIRYQLVRENNIMKIASISLV 630

RESULT 12

ADJ38132 ID ADJ38132 standard; protein; 631 AA.

XX AC ADJ38132;

XX DT 06-MAY-2004 (first entry)

XX DE Synecchococcus fn2 protein sequence.

XX KM prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;

XX KM agronomic; horticultural; crop plant; ornamental plant; woody plant;

XX KM herbicide target.

XX OS Synecchococcus sp.

XX PN MO2004001003-A2.

XX PD 31-DEC-2003.

XX PF 20-JUN-2003; 2003MO-US019536.

XX PR 20-JUN-2002; 2002US-0390140P.

XX PR 09-AUG-2002; 2002US-0402242P.

XX PR 20-JUN-2003; 2003US-00600070.

XX (UNMS) UNIV MICHIGAN STATE.

XX PI Oeteryoung KM, Viltha S, Kokscharova OA, Gao H;

XX DR WPI: 2004-082486/08.

XX DR N-PSDB; ADJ38131.

XX PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful

XX PT for further characterizing plastid division in plant cells, and in

XX PT varying agronomic and horticultural characteristics of economically

XX PT important plants.

XX PS Claim 1; SEQ ID NO 5; 287bp; English.

XX CC This invention relates to novel prokaryotic type or plastid division and

XX CC related genes and proteins. In particular, the invention relates to novel

XX CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and

XX CC compositions of the present invention are useful for further

XX CC characterizing plastid division in plant cells, in order to vary

XX CC agronomic and horticultural characteristics of economically important

XX CC plants, such as crop, ornamental and woody plants. They can also be used

XX CC as herbicide targets. The present sequence is that of a protein which is

XX CC related to the invention.

XX SQ Sequence 631 AA;

Query Match 10.9%; Score 442; DB 8; Length 631;

Best Local Similarity 24.5%; Pred. No. 3.3e-30;

Matches 189; Conservative 118; Mismatches 262; Indels 204; Gaps 30;

QY 86 IPIDPYOVIGAGQHTFLTDGIRAFBAFVSKPQFGFSDALISRRQLQAACETLSNPS 145

DB 3 IPIDYRILCVGQASADKLAEVSRDLNQSPSHFSLQAARRQLRAIAELSDPEQ 62

QY 146 RREYNE-----GLDDEBAYITDVPRWDKYPGALCVQEGGEYIVARVEALL----- 194

DB 63 RDRYDRRPFQGGLEAPSELEED--MQR1-GALLILLEGKYDRVSQLEABELLPDYDAS 119

QY 195 KERLPKSFKODVVLVVMALFLDVSRDAMALDPDPFITGYEFVEBALTLLOBEGA----- 248

DB 120 AEVYDQFARBDIALATLSSQSSIGRECRQ-----QGLEYQAAGHGRS 162

QY 249 -SSLA-----PDLRAQIDETLEETTPRYVELLGLPLGDDYAAKRLNGLSGVNITL--W 299

Db 163 QSALADHQRPPELSRTHQEQGLRPYRIIERLAQPLTAD--SDROGQLLLQAMLDNRQ 220

QY 300 SVGGGASALVGGITREKFMNEAFLR-----MTAAEQVDLF---VATPSNI PAESFEVYE 351

DB 221 GIEGPDG--SGITLTDNPL--MFLQIRGYLTILAEOQLFESBARPS--PAASF---F 271

QY 352 VALALVNAQFIQKKPHLLQDADKQFOQLQAQKNAMEIPMLYDTRNNWEIDFGLERGLC 411

DB 272 ACYTLIARGFCDHQPSLIHRASILLHELKS-----RMDVHIEQAI 312

QY 412 ALLIGKYDECRMLGLDSED-----SQYRNPAIYEVYLENSNRDNDLPLGCTLLETWLA 467

DB 313 SLILGQPEEAELLL-VQSDQDEETLSQIRALAQGEAL-----IVGLCRFTETWLA 360

QY 468 GVYFPRPRDQDKFKLGDYDDPMVLSYLERVEVVGQSPLAAATWARI GAHVKASAM 527

DB 361 TKVFPDFRDLKERTAPLQYFDDPDVQYTI DAIVEL----- 396

QY 528 QALQKVPFSRYTDRNSAPKDVQETVFSYDP-----VGNVYGRDQEPGVTAAYRPS 580

DB 397 -----PEDLMPTPLPVEPFLVRSILAKELPTPATPG-----VAPF 432

QY 581 ENFETNDYAIRAGVSESSVDETTENSVDMLKEASVKILAAVYAGLISLFSQKYLKS 640

DB 433 PRRRRRDRSERPARITAKRPLPFWIGLV-----VVVLGGGTGV-----WAMRS 475

QY 641 SSSFORKDWSSMESDVATIGSVRADDSALP-----RMDARTAEINIVSKQKIKS 691

DB 476 RSN-----STPPTPPVQVQTLPEAVPAPSPAPVVALDRAQAEVTLQNLAAKA 524

QY 692 LAFGPDHRIEMLPBVLIDGRMLKIWTDRAAETA--QLGLVYDYTLKLKSVSVTSADGTR 749

DB 525 AALGPGYDRDLATVLTGVLQWGFSSQQAQNTQLTSQFDH---KLTVDVSVLQSDGDR 581

QY 750 ALVEATLEESACSDLVHPENNAITDVRT---YTRRYEVFWSKSGKWTESVYL 799

DB 582 AVQAQKVD-----VEQYVRGDQLLETRRDGLVIRYQLVRENNIMKIASISLV 630

RESULT 13

ADJ38240 ID ADJ38240 standard; protein; 631 AA.

XX AC ADJ38240;

XX DT 06-MAY-2004 (first entry)

XX DE Plastid division-related Arc6 orthologue protein 31.

XX KM prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;

XX KM agronomic; horticultural; crop plant; ornamental plant; woody plant;

XX KM herbicide target.

XX OS Synecchococcus sp.

XX PN MO2004001003-A2.

XX PD 31-DEC-2003.

XX PF 20-JUN-2003; 2003MO-US019536.

XX PR 20-JUN-2002; 2002US-0390140P.

XX PR 09-AUG-2002; 2002US-0402242P.

XX PR 20-JUN-2003; 2003US-00600070.

XX (UNMS) UNIV MICHIGAN STATE.

XX PI Oeteryoung KM, Viltha S, Kokscharova OA, Gao H;

XX DR WPI: 2004-082486/08.

XX DR N-PSDB; ADJ38241.


```

Db      240 LARAQFLKSKVTLGKLA--LTQIIESLEGALPCTLDLGLPTPENARRGAIYAL 297
QY      295 RNLMSVGGGAGALVGLTRK-----FMNEAFLEMTAEOYDLF-----VATP 339
Db      298 RELLRQ-----GLSVEASCOIQDMPFCFLSQIISRLATEYVDLIPMDLAIYTRK 346
QY      340 SNIPASF-----VEVALALVAQAFIGKKPHLDADKQFOOLQOAKVAMEIP 390
Db      347 NKKSLSHNRVVIDNCFYMWLGHIAVGFSG-----KQNTINKAKTIGCECLI 396
QY      391 AMLYDRNMWEIDFGLEKICALLIGKVDCEKRWLGL-----DSEDQYRPALVEYLEN 446
Db      397 A-----SEGVDLKKEBAFCFLKQGSBALEKLEKQESNDSDAVRNS-----ILGK 444
QY      447 SNRDDNDLGLCKLETWLAGVVPFRPRTK-----DKKFKLGDYDDPMVL 494
Db      445 ESRSTS-----ATPSLEAMLMESVLANFPDTRGCSPLANFPRAEKYPENKMGSSSIM 499
QY      495 SY-----LERVEVQGSPLAAATWARIQAEHVKASAMQALOKVPPSRYTDRNSAPKD 548
Db      500 NHTNQRPLSTQTFVNS-----QHL-----YTAVEQLPTD 531
QY      549 VQETVPSVDVGNNGVGDGEPGVFIABAVRPSNFETNDYAIRAGVSESV-DETTVENS 607
Db      532 LQSPVVSAAK--NNDE-----TSASMPSYQLKRN-----LGVHKNKIMDEW--LS 571
QY      608 VADMLEKASVKILAAAGVAIGLISLFSQKY-----MESDVATIGSVRADSEALPR 673
Db      572 QSSLIGRVSVVAL-----LQCTVFFSLKSGINSGLQSPISVSARPHSESDSFLMKTE 626
QY      637 ---FLKSSSSFFORKDMSV-----MESDVATIGSVRADSEALPR 673
Db      627 SGNFRKNLDVNNNGIVGNIKVILDMKMGHHPALYLKSSGQATSLSHSSELHKR 686
QY      674 -MDARTNENIVSKQKIKSLAFGPDHRIEMLPYLDGRMTKITWDRAETPAQLGLV-DY 731
Db      687 PMDTEEBEELVQWENYKABALGPTHQVYSLSEVLDSEMLVOW-QTLAQTAEAKSCYMRP 745
QY      732 TLTKLSDSVTVSAD--GTRALVEATLEBSACISDLVHPENNATDVRVTTTREVFWMSK 788
Db      746 VILHLEVLQAHITBDGAGENAETLEALAEALVDESOPR-NKXYSTYKIRIYILKQOE 804
QY      789 SG-WKITEGSV 798
Db      805 DGLMKFCQSDI 815

RESULT 15
ADJ38251
ID      ADJ38251 standard; protein; 819 AA.
XX
AC      ADJ38251;
XX
DT      06-MAY-2004 (first entry)
XX
DE      Placid division-related Arc6 orthologue protein 39.
XX
KW      prokaryotic type; placid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
      agronomic; horticultural; crop plant; ornamental plant; woody plant;
      herbicide target.
XX
OS      Arabidopsis thaliana.
XX
PN      MO2004001003-A2.
XX
PD      31-DEC-2003.
XX
PF      20-JUN-2003; 2003WO-US019536.
XX
PR      20-JUN-2002; 2002US-0390140P.
      09-AUG-2002; 2002US-0402242P.
      20-JUN-2003; 2003US-00600070.
XX

```

```

PA      (UNMS ) UNIV MICHIGAN STATE.
XX
PI      Oosteryoung KM, Viltha S, Kokscharova OA, Gao H;
XX
DR      WPI; 2004-082486/08.
XX
PT      New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT      for further characterizing placid division in plant cells, and in
PT      varying agronomic and horticultural characteristics of economically
PT      important plants.
XX
PS      Disclosure; Fig 8; 287pp; English.
XX
CC      This invention relates to novel prokaryotic type or placid division and
CC      related genes and proteins. In particular, the invention relates to novel
CC      Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC      compositions of the present invention are useful for further
CC      characterizing placid division in plant cells, in order to vary
CC      agronomic and horticultural characteristics of economically important
CC      plants, such as crop, ornamental and woody plants. They can also be used
CC      as herbicide targets. The present sequence is that of a protein which is
CC      related to the invention.
XX
SQ      Sequence 819 AA;

Query Match          10.6%; Score 431; DB 8; Length 819;
Best local similarity 23.5%; Pred. No. 5e-29;
Matches 200; Conservative 116; Mismatches 311; Indels 224; Gaps 32;

QY      56 SSSSSPATATTATVLSLPSIDRPRHVPPIPDYQVLAGAQTHTLTDGIRAFPEARVSK 115
      81 SRTSSLAASTST-----LELPYTCYOLIGVSRQAEDEVKSVINLTKT 124
      116 PPOGFSDDALISRQIILQACETLSNPSRRRYNGLDDEEATVITDVPWKGALC 175
      125 DAEGTMEAAARQDLNDVDRKL--LEFSYAGNLKXIAPKBPLRPWMLBQALC 181
      176 VLOEGGETEVLAVGALTKERLPKSFQDQVVLVMAIAFDVSRDMALDPDPFITGYEF 235
      182 LLOVGQEKVLVDIGRAALRNDSKPYIHDIPLSMALBQALKAAPVNVKS--QGFEA 239
      236 VEBALKLOEE-GASSLPADLPRAQIDETLEITPRVYELGLPLDGDVYAAKGLNGISV 294
      240 LARAQFLKSKVTLGKLA--LTQIIESLEGALPCTLDLGLPTPENARRGAIYAL 297
      295 RNLMSVGGGAGALVGLTRK-----FMNEAFLEMTAEOYDLF-----VATP 339
      298 RELLRQ-----GLSVEASCOIQDMPFCFLSQIISRLATEYVDLIPMDLAIYTRK 346
      340 SNIPASF-----VEVALALVAQAFIGKKPHLDADKQFOOLQOAKVAMEIP 390
      347 NKKSLSHNRVVIDNCFYMWLGHIAVGFSG-----KQNTINKAKTIGCECLI 396
      391 AMLYDRNMWEIDFGLEKICALLIGKVDCEKRWLGL-----DSEDQYRPALVEYLEN 446
      397 A-----SEGVDLKKEBAFCFLKQGSBALEKLEKQESNDSDAVRNS-----ILGK 444
      447 SNRDDNDLGLCKLETWLAGVVPFRPRTK-----DKKFKLGDYDDPMVL 494
      445 ESRSTS-----ATPSLEAMLMESVLANFPDTRGCSPLANFPRAEKYPENKMGSSSIM 499
      495 SY-----LERVEVQGSPLAAATWARIQAEHVKASAMQALOKVPPSRYTDRNSAPKD 548
      500 NHTNQRPLSTQTFVNS-----QHL-----YTAVEQLPTD 531
      549 VQETVPSVDVGNNGVGDGEPGVFIABAVRPSNFETNDYAIRAGVSESV-DETTVENS 607
      532 LQSPVVSAAK--NNDE-----TSASMPSYQLKRN-----LGVHKNKIMDEW--LS 571
      608 VADMLEKASVKILAAAGVAIGLISLFSQKY-----MESDVATIGSVRADSEALPR 673
      572 QSSLIGRVSVVAL-----LQCTVFFSLKSGINSGLQSPISVSARPHSESDSFLMKTE 626

```



```

QY 637 ---FLKSSSFQRKDWSS-----MESDVATIGSVRADDSBALPR 673
Db 627 SGNFRKRLDSDVNRNGIVGNIKVLDMLKMHGCHPDALYKSSQSATSLSHSASBLHR 686
QY 674 -MDARTAENIVSKKQKIKSLAFSGPDHRIEMLPEVLDGRMLKIWTDRAAETAOGLVY-DY 731
Db 687 PMOTEEABEELVRQWENYKABALGPTHQVYSLSEVLDESMLVQW-QTLAQTAEAKSCYWRP 745
QY 732 TLKLSYDSDVTVSAD---GTRALVEATIESACLSDLVHPENNATDVRTYTTREYEVFWSK 788
Db 746 VLAHLEVLQAHIFPDGIAGEAAEIEALLBEAAELVDESQPK-NAKYYSTYKIRYILKKOE 804
QY 789 SG--WKITEGSV 798
Db 805 DGLMKFCOSDI 815

```

Search completed: June 10, 2005, 01:38:50
 Job time : 101 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2005, 01:02:57 ; Search time 28 Seconds
(without alignments)
2135.496 Million cell updates/sec

Title: US-10-600-070-2

Perfect score: 4063
Sequence: 1 MEALSHVIGIGLSPQLRLP.....YEVWMSKSGWKITGSLAS 801

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145.5	3.6	6095	US-09-144-085-2	Sequence 2, Appl1
2	143	3.5	2089	US-08-418-893D-23	Sequence 23, Appl1
3	143	3.5	2089	US-08-418-893D-24	Sequence 24, Appl1
4	139.5	3.4	954	US-09-057-969-2	Sequence 2, Appl1
5	131.5	3.2	1194	US-08-680-326-35	Sequence 35, Appl1
6	128.5	3.2	876	US-08-436-664-20	Sequence 20, Appl1
7	128.5	3.2	876	US-08-436-664-32	Sequence 32, Appl1
8	128.5	3.2	876	US-08-436-664-34	Sequence 34, Appl1
9	128.5	3.2	876	US-09-135-642-20	Sequence 20, Appl1
10	128.5	3.2	876	US-09-135-642-32	Sequence 32, Appl1
11	128.5	3.2	876	US-09-135-642-34	Sequence 34, Appl1
12	128.5	3.2	876	US-08-394-232A-22	Sequence 20, Appl1
13	128.5	3.2	876	US-08-394-232A-32	Sequence 32, Appl1
14	128.5	3.2	876	US-08-394-232A-34	Sequence 34, Appl1
15	128.5	3.2	876	PCT-US95-04080-20	Sequence 20, Appl1
16	128.5	3.2	876	PCT-US95-04080-32	Sequence 32, Appl1
17	128.5	3.2	876	PCT-US95-04080-34	Sequence 34, Appl1
18	128.5	3.1	2482	US-09-252-991A-16967	Sequence 16967, A
19	125.5	3.1	5087	US-09-144-085-1	Sequence 1, Appl1
20	122.5	3.0	2756	US-08-375-709-11	Sequence 11, Appl1
21	122.5	3.0	2756	US-08-753-939-11	Sequence 11, Appl1
22	122.5	3.0	2756	US-09-090-793-7	Sequence 7, Appl1
23	122.5	3.0	2756	US-09-231-899-7	Sequence 7, Appl1
24	121.5	3.0	680	US-09-252-991A-26639	Sequence 26639, A
25	121.5	3.0	947	US-09-540-236-1991	Sequence 1991, Ap
26	121.5	3.0	1220	US-09-540-236-3011	Sequence 3011, Ap
27	120	3.0	852	US-09-585-858-19	Sequence 19, Appl1

28	120	3.0	852	4	US-10-270-878-19	Sequence 19, Appl1
29	120	3.0	1262	4	US-09-198-452A-97	Sequence 97, Appl1
30	120	3.0	1262	4	US-09-438-185A-83	Sequence 83, Appl1
31	119.5	2.9	831	1	US-08-073-384C-5	Sequence 5, Appl1
32	119.5	2.9	831	1	US-08-254-359A-5	Sequence 5, Appl1
33	119.5	2.9	831	1	US-08-463-043-5	Sequence 5, Appl1
34	119.5	2.9	831	1	US-08-481-238-5	Sequence 5, Appl1
35	119.5	2.9	831	2	US-08-471-066B-5	Sequence 5, Appl1
36	119.5	2.9	831	2	US-08-484-956-5	Sequence 5, Appl1
37	119.5	2.9	831	2	US-08-757-653-5	Sequence 5, Appl1
38	119.5	2.9	831	2	US-08-599-491-5	Sequence 5, Appl1
39	119.5	2.9	831	2	US-08-756-386-5	Sequence 5, Appl1
40	119.5	2.9	831	2	US-08-823-516-5	Sequence 5, Appl1
41	119.5	2.9	831	3	US-08-682-853A-5	Sequence 5, Appl1
42	119.5	2.9	831	3	US-08-759-038-5	Sequence 5, Appl1
43	119.5	2.9	831	3	US-08-758-314-5	Sequence 5, Appl1
44	119.5	2.9	831	3	US-09-350-309-5	Sequence 5, Appl1
45	119.5	2.9	831	3	US-08-520-946-5	Sequence 5, Appl1

ALIGNMENTS

```
RESULT 1
US-09-144-085-2
; Sequence 2, Application US/09144085
; Patent No. 6280999
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Claes
; APPLICANT: Betlach, Mary C.
; APPLICANT: Ashley, Gary
; APPLICANT: Julien, Bryan
; APPLICANT: Ziemann, Rainer
; TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
; FILE REFERENCE: 30062-20020.20
; CURRENT APPLICATION NUMBER: US/09/144, 085
; EARLIER FILING DATE: 1998-08-31
; EARLIER APPLICATION NUMBER: 09/010, 809
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 6095
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-144-085-2

Query Match          3.6%; Score 145.5; DB 3; Length 6095;
Best Local Similarity 19.3%; Pred. No. 0.0034;
Matches 167; Conservative 128; Mismatches 349; Indels 221; Gaps 38;

70  LVSLPPSIDPREERA-VPIPIFYVLGAQTHFLTDGIRRAFEARVSKPPQGFSDALIS 128
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1642 LALLDALSRREALIVVHLDIAQ-----LQRLGSSGELPALF-----RALTL 1684

129 RKQLQAACETLSNPKRSRREYNEGILDDDEKATVTVDPMDKVPALCYLQGGETEYLR 188
    :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
1685 RPSLRKKSASATRRPASLRERLSALPEARLNALVEVRGEV-AAVAGLQ----- 1734

189 VGEALNERLPKSKQKQVVLVMAALAFDVSRDMAALPPDPITGYEFVEALKLQEEGA 248
    |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
1735 GEAVALADQVKEIGLQSLMAVLAIRNRLTSTKSTSL-PATIVFYPTPRRAIELLKQAF 1792

249 SSL-APLRAQIDETLEITPRVYLEL-----GLPLGDDYAAKRLNGLSGVNRI--LWS 300
    ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
1793 SGLQVKEARARVRRRAKDEPIALVSMACRLPGVAIAPPDYWRLLAGKQALIEGLPARWD 1852

301 -----VGGGASALVGGLTREKPMNER-----FLKMTAAQVVD 333
    |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
1853 GFEEYDDPEAKASVYAREGGFVADIDLFDANFGISPREASMDPOHRLVLETAWALE 1912
    ||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
334 LFVATPENIPAESFEYVVALA-----LVAQAFTIGKKPHLLQDADKQFOQLQ 380
```

```

Db      1913 RAGVPSALSGSAGVYIGSMGSDYGALHTVLDLKELEGYRGISGAASITSG-----R 1964
Qy      381 QAKVMAEITPAMLYDTNNMNEIDFGELRGICALLIGKVDECRMWIGLDSBDSQYRNPAI- 439
Db      1965 VAYALGIQGPMTYDVYICSSSL-VSLHACTALNQECB-----LALAGVTWSTPLAF 2018
Qy      440 VEFV-LENSNRDNDLPGCLKLET-----WLAG---VVFPRFRDTKCKKFLGDIYD 489
Db      2019 VEFSLKGMGRD-----GRCKSPSVQADAGMAGCGMILLKTLSDAQ-----RD 2063
Qy      490 DPMVLSL-----EKEVYVQ-----504
Db      2064 GDRVGLVIRGSAAVQDGRSQGLTPANGPAQQRVIRQALSSGSLSPEDIDAVEAHGTGSL 2123
Qy      505 GSPFLAAATATARGAE-----HYKASAMQA---LOKVPFSRYTDRNSAE 545
Db      2124 GDTIEGALAEVGFEPSPERPLYLGSSSKNLGHAQAAAGVAYIKVYL-----SMQ 2175
Qy      546 PKDQETVFSVDPVGNVNGRDEBPVFIAEAVRPSBNFETNDYAIRAGVSSSVDETVE 605
Db      2176 HEVLPKTLHAEP-SPHIGWGS-GLSLQEARP---WRNNGRVRAVGSVFGISGTNAH 2230
Qy      606 MSVADMLKEA---SVKLLAGVAIGLISPSQKFLKSSSSFOKDMVSSME---SDYAT 659
Db      2231 IILEEAPAEARRBEVEEAAPALLPLVLSGRDEAAVNAQAQRMKMLEHGEVGSMDVVR 2290
Qy      660 IGSVRAADSALPRMDARTAEINIVSKWQIKSLAFG-PDHRILEMPLVLDGRMLKIWTR 718
Db      2291 TAAHHRHFSRSRAVLAASAGAV---EGLRALSSGRPDAAVSGTAKRGKLAVALTTGQ 2347
Qy      719 AETAQAG--LVYDYTLKLKLSVDSVTVSAD-----GTRALVEATL--EESACT--SDLVH 767
Db      2348 GSGRLGKGLKLYEYVYFRAAFDEVCEALDAYLRGRLREVFAAAGEEGALLERTETYG 2407
Qy      768 PENNAIDVYRTTYRTYRVEVFNKSGMK 792
Db      2408 PGLFALEVALYRQ-----WESWGLK 2427

RESULT 2
US-08-418-893D-23
Sequence 23, Application US/08418893D
Patent No. 5559220
GENERAL INFORMATION:
APPLICANT: ROESSLER, PAUL G
APPLICANT: OHLROGE, JOHN B
TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A
TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTHELLA CRYPTICA
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY
STREET: 1617 COLE BLVD.
CITY: Golden
STATE: CO
COUNTRY: USA
ZIP: 80401-3393
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/418,893D
FILING DATE: April 7, 1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/104,938
FILING DATE: September 14, 1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: O'CONNOR, EDNA
REGISTRATION NUMBER: 29,252

```

```

; REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303-231-1000
; TELEFAX: 303-231-1098
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2089 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-418-893D-23

Query Match      3.54; Score 143; DB 1; Length 2089;
Best Local Similarity 20.04; Pred. No. 0.0009;
Matches 169; Conservative 113; Mismatches 289; Indels 272; Gaps 44;

Qy      73 LPSIDRPERHVPIDIFYQVLAQTHFLTDGIRRAFEARVSKPP-----QFGFSDAL 126
Db      197 LFNALDK-----LGIFIGPTGPMVSLGDKIAANILAQTAVPSPWGSFGGPDGP 250
Qy      127 ISRRQIIQAACETLSNPRSREYNEGLLD--DEBAYITDVPMDKPGALCVLQEG-- 181
Db      251 LQ-----ADLTBEETGTPMEIFNKGGLVTSADEAVIYANKIGWEN--GIMIKASEGGGK 301
Qy      182 -----ETREIYLR-----VGEALLKERLPKSPQDVVLWALAFLVSRDAMALDP 226
Db      302 GIRFVDEADLRNAFVQVSNVEIGSPIFLMQLCKNAHIEVQIVG---DQGNANALNG 357
Qy      227 PDFITGEYFEVEALKLQEBGASLAPDLRAQIDETLEETIPRYVELGLPLGDDVYAK 286
Db      358 RDCSTGRF-----QKIFEEGPPSIYP-----KETFHEK-----ELAAQ 391
Qy      287 RLNLGSGVRNIIILSVGGGASALVGLTRKFPNMEAFLRMTAAEQVDLFVA---TSPNIP 343
Db      392 RL-----TONIGYO--GAGTVEYLYNNAADNKF--FLEINPLQYEHPTBEGITGNLP 441
Qy      344 AES-----PEYVEVALALVAQAFIGKKPHLDQADKQFOLOQAQVMAEIPA--- 391
Db      442 ATQLQVAMGIPFNIPDIRLYGREDAVGTDP--IDFLQRYEL--DSHVIYARITENP 498
Qy      392 -----MYDTRNNW-EIDFGELRGICALLIGKVDECRMWIGLDSBDSQY-- 434
Db      499 DEBPKPTSSGIERIKQGSTPNWGVFSVGANGI-----HEFADSGFGH 542
Qy      435 ---RNP-----AIVEFLSNSRDD-NDLPGCLKLET-----WLAGVVP 472
Db      543 LPAKGPMBQARRKALVIALKEMEVRGDIRNSVEYLVKLETTFAFKNTIDTWSLDGII-- 600
Qy      473 RFRDIDKPKLDDYDDPMVLSYLEKVEVVOGSPPLAAATMARIGAEHFKASMQA--- 529
Db      601 ---KEKSVKV-----EMPSHL---VVGAAGFAKAEHVKVATBEVKSFRKQVS 644
Qy      530 -----LOKVPFSRYTDRNSAEPKDQET-----552
Db      645 TAGIPGINSFINIVAYIDTKYPRHVERISPDVYRFLTDGTTIVVYQTALGALLATFGG 704
Qy      553 ---VPSVD-PVGNVNGRGE---PGVFIAEVRBS-----ENFETNDYAIRAGVSES 597
Db      705 ETRHIFGMDPEPLRLSLDQATVLMPTIEPSELRLDVTGKVVRYLQDNGATYEAQ--- 760
Qy      598 SVDETTVMSVADMLKEASVKIILAAGVAILGLISLFSQKYFLKSSSFQKDMVSSME-SD 656
Db      761 ---QPYVEVAMKMT--MPTKATESG-----KITHNLISGSVTSAGDLASLELKD 806
Qy      657 VATIGSV-----RADSEALPRADARTA--ENIVSKWQIKSLAFGPDHRIEMPEVLDR 710
Db      807 PSRVKKIETFSGLDDMESKVDLEPKAVMNVLS-----GFNLD-----PEANAQ 852
Qy      711 MKIWTDRAAETAQLGLVYD-VYTLKLKLSVDSVTVSADGTRALVEATLEESACLSDLVHP 769

```

Db 853 AIDSATSSAAADLLVQVLDDEFYVESQFDGV-IADVVRLTLTANTE-----TLDVIVSE 907

Qy 770 NNA 772

Db 908 NLA 910

RESULT 3

US-08-418-893D-24

/ Sequence 24, Application US/08418893D

/ Patent No. 5559220

/ GENERAL INFORMATION:

/ APPLICANT: ROESSLER, PAUL G

/ APPLICANT: OHLROGE, JOHN B

/ TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A

/ TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTELLA CRYPTICA

/ NUMBER OF SEQUENCES: 25

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: NATIONAL RENOVABLE ENERGY LABORATORY

/ STREET: 1617 Cole Blvd.

/ CITY: Golden

/ STATE: CO

/ COUNTRY: USA

/ ZIP: 80401-3393

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/418,893D

/ FILING DATE: April 7, 1995

/ CLASSIFICATION: 800

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 08/104,938

/ FILING DATE: September 14, 1993

/ CLASSIFICATION: 800

/ ATTORNEY/AGENT INFORMATION:

/ NAME: O'CONNOR, EDNA

/ REGISTRATION NUMBER: 29,252

/ REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 303-231-1000

/ TELEFAX: 303-231-1098

/ TELEX:

/ INFORMATION FOR SEQ ID NO: 24:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 2089 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULAR TYPE: protein

/ HYPOTHETICAL: NO

/ ANTI-SENSE: NO

/ FRAGMENT TYPE: N-terminal

/ US-08-418-893D-24

Query Match 3.5%; Score 143; DB 1; Length 2089;

Best local Similarity 20.0%; Pred. No. 0.0009;

Matches 169; Conservative 113; Mismatches 269; Indels 272; Gaps 44;

Qy 73 LPSIDRPERHVPIDFYQVLAQTHFLNDGIRRAFEARVSKP-----QGFPSDAL 126

Db 197 LFNALDK-----IGIKFIPGTGPMVSVLGDKIANITLAQAKVPSIPWGSFGPDGDP 250

Qy 127 ISRQIILQAACETLSNRSRREYNEGLD--DEBAYITVDPMKVPKALCVLQEG--- 181

Db 251 LQ-----ADLTEEGTIPMEIFPNKGLVTSADNAVYANKIKGWN--GIMTKASEGGGCK 301

Qy 182 -----ETREIVLR-----VGEALIKERLPSFKQDVVLAVALFLVDSRDAMALDP 226

Db 302 GIRFVDHEADLRNMFVQVSNVEVIGSPIFLMQLCKNAHRIEIVQIG---DQHGNAVALNG 357

Qy 227 PDFITGEFVEALKLQEBGASLADPLRAQIDETLEETPRRYVLELGLPLGDDYAAK 286

Db 358 RDCSTORRF-----QKIFEEGPESTIV-----KETHHEM-----ELAAQ 391

Qy 287 RLNLGSGVRNLSVGGGASALVGLTREKFNNEAFLRMTAAQVLDLFAA---TPSNIP 343

Db 392 RL-----TONIGVQ--GAGTVEYLYNADNKF--FLEINPRLOVHEHPVTEGITGANLP 441

Qy 344 AES-----FEYVEVALALVAQAFIGKKPHLLQDADQFQOLQQAQKMMAMEIPA--- 391

Db 442 ATOLQVAMGIPLEFNIPIRLRYGREDAYGDP--IDFLQERVEL--DSHVIABITLENP 498

Qy 392 -----MLYTRNNW--EIDFGLERGCALLIGKVCRCMWLIGDSBSQY-- 434

Db 499 DEGRKPTSGSIEKFKPOSTRWVWGFVSGANGI-----HEFADSQGH 542

Qy 435 ---RNP-----AIVEFLSNSRDD--NDLPGLCKLET-----WLAGVFP 472

Db 543 LFAKGPMBQARKALVIALKEMEVRGDIKNSVEYLVKLTETFAFKKNTIDTSMLDGIT-- 600

Qy 473 RFBDTKKKRKLGDYDDPMVLSYLERVEVYQSPPLAAATMAIGAEHVYASAMQA--- 529

Db 601 ---KERSVKV-----EMPSHL-----VVVGAIVKAEFEHVAVATEEVESFRKQVVS 644

Qy 530 -----LOKVPSPRYTDRSABPKVOET----- 552

Db 645 TAGIPGINSRNIEVAYIDTKPFVEVERISPDVYRFTLDGNTIDVEVYOTAGALLATTGG 704

Qy 553 ---VPSVD--PVGNNGRDBE---PGVFLAEAVRPS-----ENFTNDYAIRAGVSES 597

Db 705 ETRHIFGMDEBLGRLSLDGAATVLMPTIFDPSBLRTVTKGVVYLDNGATVAG--- 760

Qy 598 SVDETTVMGVADMLKASVYKILAGVAILGISLPSQKTFKSSSSFORQDMVSME--SD 656

Db 761 ---QPYVEVEAMKMI--MPIKATESG-----KITHNLSAGSVISAGDLLSLLEKID 806

Qy 657 VATIGSV-----RADSEALPRMDART--ENIVSKWKIKSLAGPDRHLEPEVLDGR 710

Db 807 PSRVKIKETTSKIDINESKYVDLEPQAVMNVLS-----GNLD-----PEAVAAQ 852

Qy 711 MLKIWTDRAAETAGLGVYD--YTLIKLVSVDVTSADGTRALVEATLESACLSDLVHPE 769

Db 853 AIDSATSSAAADLLVQVLDDEFYVESQFDGV-IADVVRLTLTANTE-----TLDVIVSE 907

Qy 770 NNA 772

Db 908 NLA 910

RESULT 4

US-09-057-969-2

/ Sequence 2, Application US/09057969

/ Patent No. 6013451

/ GENERAL INFORMATION:

/ APPLICANT: WONG, VICTOR THI WONG

/ APPLICANT: PHANG, SENG MENG

/ TITLE OF INVENTION: BACILLUS STEAROTHERMOPHILUS DNA

/ TITLE OF INVENTION: POLYMERASE I (KLENOW) CLONES INCLUDING THOSE REDUCED TO

/ NUMBER OF SEQUENCES: 17

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: FROMMER LAWRENCE & HAUG LLP

/ STREET: 745 FIFTH AVENUE

/ CITY: NEW YORK

/ STATE: NEW YORK

/ COUNTRY: USA

/ ZIP: 10151

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/057,969
 FILING DATE: 09-APR-1998
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: KOMALSKI, THOMAS J.
 REGISTRATION NUMBER: 32,147
 REFERENCE/DOCKET NUMBER: 674507-2001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-588-0800
 TELEFAX: 212-588-0500
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 954 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: amino acid
 US-09-057-969-2

Query Match 3.4%; Score 139.5; DB 3; Length 954;
 Best Local Similarity 22.0%; Pred. No. 0.0005;

Matches 175; Conservative 132; Mismatches 315; Indels 173; Gaps 44;

99 HPLTDGIRRAFEARVSKPQFGFSDALISRQILQACETLSNPSRREY-NEGLDDE 157
 183 HYEADDIIGMAARAERE--GFAVKYISGDRDLTQ-----LASPQVTEITKKGITDIE 234
 158 EAVYITDY-----PMDKVPG-----ALCVQOEGGETIVYLRVG 190
 235 SYPTETVEYKGLTPEQIVDLKGLMGDKSDNIPGVPGIKKTAAVKLKQETGVNVLASI 294
 191 EALKERLPKSPKO--DVVLVMAALPDLVSRDA--MALPDPDITGYEFEEBAKLQOE 245
 295 DEIKGEKCKENLRKYRDLAL-LSKQLAICRDAAVEELTD--DIVYGEDEBKVVALFOE 351
 246 EGASSLAPDLRAQIDE-----TLEETPRVYLELGLT--PLGDDYAAKRLNG- 290
 352 LGFQSFIDKMAVQDEGEKPLAGMDFAIDSVTDEMLADKAAVLVEVVGNDYHNAPIVGI 411
 291 -LSGVRNII--SYGGGASALVGGTLREKFM-----NEAFLEMTAAEQV-D 333
 412 ALNERGRFRLPETAADPKFLMLGDETCKTMTFDSKRAVALNKGIEIAGVGVFD 471
 334 LFVATPSNIPAES-----FEVVEVALVAQAFIGK-KPHLLDADKQFOOL--QOAK 383
 472 LLLAAVILLDPAQAAGDVAANAQKMQYEAVNSDEAVYGGAKRTVPDSEPTLAEOLVKRAAA 531
 384 VMAMEIPAMLYDTRNNME-IDFGLEKGLCALLIG-----KVDECRM-WLGLD--SEDSQY 434
 532 IMALEEFMLDELRNEDRLLTELEHNLAGILANMEFTGVVDTKRLQEGAGALTEQLQ- 590
 435 RNPALVEFVLENSNRDNDLPGCLKLETLVLAGVPPRRFDITOKKPKGLGDIYDDPMVL 494
 591 ---AVERRIYELGQEFNINSP--KQLGT---VLDPKQLPVLKKTIKTG--VSTSAVL 639
 495 SYL--ERREVVQGSPLAAMATMARIAGAEVYKASAMQALQKVPSPRYTORNSAPROVET- 552
 640 EKLAIPHHEIYE--HILHYRQIGKQSTYIE-GLLKVVHPVTGKVTMFOA---LQYTG 692
 553 -VESVDPVGNVNGDGPVFIAEAVRPS-----NEFTNDYAIRAGVSESSVDETT 603
 693 RLSSVEBNLQNIPIRLBEGKRIRQAFVPSBDMILFADYSQIEIARVLAHIAE---DDNL 749
 604 VE-----NSVADMLKE--ASVKIILAAGVAILGISLFSQKTFYKSSSSSQRK 647
 750 IEAFRRWLIDHTKTAMDIFHYSEEDVTANMRQAKANFGLVIGISD--YGLAQMLNITRK 808
 648 DMSVMSDVAAITGVSADSEALPRMDARTAEIVKWKIKSLAGDPDRIMLEPEVL 707
 809 EAABFIERIYASFPGVK-----QYMDNIVQE-AKQKGYVTTLHRRRYLPDIT 855

QY 708 DGM-LKIWDRAA-ETAOGLVYDYTLKLSVD-SYTVSADGTRA-----LVEA 754
 DB 856 SRNFVTRFAERATMTYPIQGSAAD-IKKAMWIDLSVSVREERLQARLLQGHDELIEA 914
 QY 755 TLEESACLSLVHPE 769
 DB 915 PKEIRGLCRLV-PE 928

RESULT 5

US-08-680-326-35
 Sequence 35, Application US/08680326
 Patent No. 5925733
 GENERAL INFORMATION:
 APPLICANT: ROSE, TIMOTHY M.
 APPLICANT: BOSCH, MARINIX
 APPLICANT: STRAND, KURT
 APPLICANT: TODARO, GEORGE J.
 TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES
 TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL
 NUMBER OF SEQUENCES: 152
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 755 Page Mill Road
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1018
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/680,326
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Schiff, J. Michael
 REGISTRATION NUMBER: 40,253
 REFERENCE/DOCKET NUMBER: 29938-20001.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 813-5600
 TELEFAX: (415) 494-0792
 TELEX: 706141
 INFORMATION FOR SEQ ID NO: 35:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1194 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-680-326-35

Query Match 3.2%; Score 131.5; DB 2; Length 1194;
 Best Local Similarity 19.0%; Pred. No. 0.0046;
 Matches 168; Conservative 122; Mismatches 297; Indels 297; Gaps 44;

QY 48 SDNFPSDSS-----SSPATTTATIVSLPESIDRPERHY--PIPIDFYOVLAGQ--TH 99
 DB 205 NDATLNGDKNAFPGTSKASPSFREV--TERIDVYVYDTQPCAFYRVYBSSSFTN 261
 QY 100 FLTDGIR---RAFEARVSKPQFGFSDALIS-----RROQI--- 132
 DB 262 YLCDNPHPELAKXKEGRVDAATTRFLMNPGRFVSGWYQKRGVDGERVRVAPASRQLTSD 321
 QY 133 LQAAC-----ETLSNPSRREYNEGLIDE-----EAYITDVPMDKVPKALC 175
 DB 322 VEIDCMSDNLQAIIPNDSDWPDYKLLCFDIECKSGSGSNELAPDPATHLEDL---VIOISC 377
 QY 176 VL-----QEGGETIVYLRVGGALKERLPKSPKO-----DVLVMAALPDLVSRD 220
 DB 378 LLYSIPQSLHILLPSLSSCDLPORYQEMKAGLDEPTVLEPDEFFELLIAFMILVKQ 437


```

QY 221 AMALDPDPTTGYEVEEALKLOEGASIASAPLRAQIDETLEETPRVYLIGLPLG 280
D 438 Y-----AEFATGVNIIVPDAFI-----MEKINSYSLKI----- 468
QY 281 DDYAAKRLNGISGVRLNLSVGGGASALVGLTRFKFMEAFIAPRAEVDLFPATPS 340
D 469 DGYSIRNGGLFKI-----MDVKGSG-----FORSKVKINGLISLDMVAALATE 512
QY 341 NIPASEVEVEVA-----LATVAQAFIG 363
D 513 KKLKSSYKLSVAEALNESKRDLPYKDIPOGYASGPNTRGIIIGBYCIQDSALVGLKFPK 572
QY 364 KKHLLQDADQKQFOOLQOAKYAMEIPAMLDTNNMEIDPGLERGLCALLIGKVBECRM 423
D 573 YLPHLELSA-----VARLARITLTALYD-----GQVRIYTCGLGLASS--- 612
QY 424 WLGLDSDSOYRNDPAIYEP-----VLENSNRDNDLPGCLKLETWLAGV-VFPR 473
D 613 -RGFILPDGKY--PATFEYKQVLPDVGDEEEMDESDVSPGTSSGRNVGKARVFD- 668
QY 474 FPDTKKKFKLGDYDDPMVLSYERV-EVYQSSPLAAATMARI GAHVYKASAMOLQK 532
D 669 -PDT-----GFYIDPVVYLDFASLVPSIIQAHL--CFTTLTLPETVK-----R 710
QY 533 VPFERYTDRNSAEPRQVQETVFSYDP-----VGNNVGRDGEPRGFIAE--AVRSENEP 585
D 711 LNPEDY-----ATFTVGGKRLFFVRNSV-RESLGLVLLKMDLMLMK----- 750
QY 586 NDVAIRAGVSESVDETTVENSVA DMLKEASVKILA-----AGVAIGLIS----- 630
D 751 -AIRARIPSSSDE-----AVLLDKQAAIKVVCNSVGFYGAQFPLCLVVAATVT 802
QY 631 -----LFSQKYRLKSSSFQR-----KDMVSSMESVA-TIGSVARDBEALPRDAR 677
D 803 TIGROMLSTRDYIHNWMAAFERITAPPDIESSVLSQKAYEVKVIYGDTSVFIRFGV 862
QY 678 TAENIVSKMOK-----IKSLAFGPDHRIEMLPVLDGRMLKIWTRAAETAOLGLVYDTL 733
D 863 SVEGIAIKERMAHIIITALEFCPFIKLECEKTFI--KLLI-----TKKKIYIYGGKV 915
QY 734 LKLVDSDVTS-----ADGTRALVEATL-----EESACLS 763
D 916 LMGVLDLVKRNKQCFINDYARKLVELLLYDDTVSRAAAEASCVS 959

```

```

RESULT 6
US-08-436-664-20
; Sequence 20, Application US/0843664
; Patent No. 5874282
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATTHOOR
; APPLICANT: TUDOR, STARLA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; TITLE OF INVENTION: STEAROTHERMOPHILUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: PaateSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,664
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 536

```

```

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/394,232
; FILING DATE:
; APPLICATION NUMBER: 08/307,410
; FILING DATE: 16-SEP-1994
; APPLICATION NUMBER: 08/222,612
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fisher, Carlos A
; REGISTRATION NUMBER: 36,510
; REFERENCE/DOCKET NUMBER: GP94003.CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-7929
; TELEFAX: 619-546-7929
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 876 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; US-08-436-664-20

Query Match 3.2%; Score 128.5; DB 2; Length 876;
Best Local Similarity 21.0%; Pred. No. 0.0053;
Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

QY 99 HFLTDGIRBAFEAVSKPQGFSDNALISRRQILQACETLSNPRSREX-NEGILDE 157
D 107 HYENDDIIGTMAABABE--GFAVKYISGRDLTQ-----LASPVYVELTKKGIDIE 158
QY 158 EATVITDV-----PMDKVPG-----ALCVLQEGGETEIVLRVG 190
D 159 SYTETVEVVEKGLPQEIQVLDKGLMGDKSDNIPGVPIGKETAIVKLLKQFTVENVLASI 218
QY 191 EALKERLPSFKQ--DVLVLMALAFDVSQDA--MALDPDPTTGYEVEEALKLOE 245
D 219 DEIKGEKIKENTLRQYRDLAL-LSKQLAAICRDAVEVELTD--DIVYGEDEKXVALAFOE 275
QY 246 EGASSLAPDLRAQIDE-----TLEETPRVYLIGL-----PLGDDYAAKRLNG- 290
D 276 LGFQSFIDKNAVQTBEGEKLPLAGMDFAIADSVTDEMLADKALVVEVGVGDYHNAPIVGI 335
QY 291 -----LSGVNIIILSVGGGASALVGLTRFKFMEAFIAPRAE----- 330
D 336 ALANERGRFLRPRFTALADPKFLAW-----LDETKKTKMFPSSKRAVALKWKGI 385
QY 331 -----QVDLPVATPSNIPAS-----FEYEVALLVAQAFIGK-----KPHILQ 370
D 386 ELRGVDFDLILAAIYLDPAQAAGDVAAVAKKHQYEAVSDEAAVVGKAKRTVPDEPTLAE 445
QY 371 DADQFOOLQOAKYAMEIPAMLDTNNME-IDFGLERGLCALLIG-----KYDECRM 423
D 446 HLAH-----KAAAIWALBEPMLDELIRNEQDRLLTLEBQPLAGILANMEFPGVKVDTRL 500
QY 424 -WLGLD-SEDSQYRNPAIVEFVLENSNRDNDLPGCLKLETWLAGVFPFRPDTXDK 481
D 501 EQMGAEILTEQIQ-----AVERRIYELAQEFYINP-----KQIGT-----VLPKQLPLVKK 549
QY 482 PKLGDYDDPMVLSY-ERVEVQSSPLAAATMARI GAHVYKASAMOLQKVPSPRYTD 540
D 550 TKTG-YSTADVLEKRLAPHHEIYV--HILHYRQLGKLQSTYIB-GLKVHAPVYGVKHYTM 605
QY 541 RNSAEPDQVET--VFQVDPVGNVGRDGEPRGFIAEVRSE-----NFTNDPAI 590
D 606 FNOA-----LTQTGRLSSVEPALONIPIRLERGRKIRQAFVSESDMLIFADYISQIEKRV 661
QY 591 RAGVSESSVDTEVE-----MSVADMLKE--ASVKILAAGVAILGILSFSQ 634

```

```

Db      662 LAHIAE---DNLIEAFRRGLDHTKTAMDIFHVSEEDVTANMRQAQAVNFGIVGISD 718
      635 KYFLKSSSSFORQDMVSMESDVATIGSVRADSEALPRMDARTANIVSKWQIKSLAF 694
      719 -YGLAQNLTNTRKEAAFLIERFYFASFPQVK-----QYMDNIVQE-AKQKGYVT 764
Qy      695 GPDHRIEMLEPEVLDGRN-LKIWTDRAA-ETAOQLGVYDYLTKLSVD-SYTVSADGTRA- 750
Db      765 TLHRRRYLPDITSRRNFVNSFAERTAMNTPIQGSAD-IKKAMIDLSTRLAEERLQAR 823
Qy      751 -----LVEATLESACLSDLVHPE 769
Db      824 LLLQVHDELILEAPKEIERLCRLV-PE 850

```

RESULT 7

```

US-08-436-664-32
; Sequence 32, Application US/08436664
; Patent No. 5874282
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATHOOR
; APPLICANT: TUDOR, STARLA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,664
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/394,232
; FILING DATE:
; APPLICATION NUMBER: 08/307,410
; FILING DATE: 16-SEP-1994
; APPLICATION NUMBER: 08/222,612
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleher, Carlos A
; REGISTRATION NUMBER: 36,510
; REFERENCE/DOCKET NUMBER: GP94003.CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-2807
; TELEFAX: 619-546-7929
; TELEX:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 876 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-436-664-32

```

```

Query Match      3.2%; Score 128.5; DB 2; Length 876;
Best Local Similarity 21.0%; Pred. No. 0.0053;

```

```

Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;
Qy      99 HPLTDGIRAPFAVRVSKPQFGPSDDALISRRQILQAACETLSPRRREY-NEGILLDE 157
      107 HYEAADDIGTMAARERE---GFAVKYISGDRLTQ-----LASPQVTVEITKKGIIDIE 158
Qy      158 EATVITDV-----PMDKVG-----ALCYLQEGGEIYILRVG 190
Db      159 SYPTETVEKYGTLPEQIVDLKGLMGDKSDNITGVGQIGKTVKLLKQPGTEVENVLASI 218
Qy      191 BALLKERLPKSPFO--DVVLWALAFLDVRDA---MADPPFITGYEEVEALKLLQE 245
Db      219 DEIKGEKLENLQRYDLAL-LSKOLAAICRDPVETLTD--DIYKGEERKQVALLFOE 275
Qy      246 EGASLAPDLRAQIDE-----TLEETIRPVYELIAGT---PLGDDYAKKLNG- 290
Db      276 LGFQSLDKMAVQTDGSEKPLAGMDFAIDSVTDEMADKALVVEVGNVHNAPIVGI 335
Qy      291 -----LSCVRNLIKSVGGGASALVGLTREFKFNDEALFMTAAE---- 330
Db      336 ALANERGFELRPETALADPKFLAM-----LGDETKKKTMDPSKRAVALKWKGI 385
Qy      331 -----QVDLEFVATPSNIPAES-----PEVVEVALALVNOAFITG-----KPHLQ 370
Db      386 ELRGVVFDELALAYLLDPAQAGDVAAVAMQGYEAVRSDEAVYGGAKRTVDEEPTLAE 445
Qy      371 DADKQFOQLQOAKVMMEIPAMLDTNNME-IDFGLERGLCALLIG-----KVDECRM 423
Db      446 HLAR-----KAAAIWMLBEPMLDELNRNEQDRLLTELEQPLAGILANMEFTGVKVDPKRL 500
Qy      424 -WIGLD-SEDSQYRNPAIYEFVLENSNRDNDLPGICLLETWLAGVVEPRRDYDKK 481
Db      501 EOMGAEITEQLQ-----AVERRIYELAQEFNINSP--KOLGT-----VLPDKLQLPVTKK 549
Qy      482 FXIGDYVDDPMVLSYL-EREVEVQSGSPLAAATMAATIGAEHVASAMQALQKVPSPRYTD 540
Db      550 TKTG-YTSDADVLEKLAIPHHEIVE--HILHYROLQKIQSYIE-GLUKVHPVYTKVHTM 605
Qy      541 RNSAEPDVOET--VFSVDPVGNVGRDGEFVFIKAVRPSB-----NFEINDYAI 590
Db      606 FNGA-----LTQTRLSSVEENLQNIPIRLEEGKIKQAFPSPDMLFPAADYSQLELRV 661
Qy      591 RAGVSSSVDETTVE-----MSVADMLKE--ASVKIILAAGVALGLISFSQ 634
Db      662 LAHIAE---DNLIEAFRRGLDHTKTAMDIFHVSEEDVTANMRQAQAVNFGIVGISD 718
Qy      635 KYFLKSSSSFORQDMVSMESDVATIGSVRADSEALPRMDARTANIVSKWQIKSLAF 694
Db      719 -YGLAQNLTNTRKEAAFLIERFYFASFPQVK-----QYMDNIVQE-AKQKGYVT 764
Qy      695 GPDHRIEMLEPEVLDGRN-LKIWTDRAA-ETAOQLGVYDYLTKLSVD-SYTVSADGTRA- 750
Db      765 TLHRRRYLPDITSRRNFVNSFAERTAMNTPIQGSAD-IKKAMIDLSTRLAEERLQAR 823
Qy      751 -----LVEATLESACLSDLVHPE 769
Db      824 LLLQVHDELILEAPKEIERLCRLV-PE 850

```

RESULT 8

```

US-08-436-664-34
; Sequence 34, Application US/08436664
; Patent No. 5874282
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATHOOR
; APPLICANT: TUDOR, STARLA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive

```

CITY: San Diego
 STATE: CA
 COUNTRY: USA
 ZIP: 92121
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/436,664
 FILING DATE: 08-MAY-1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/394,232
 FILING DATE:
 APPLICATION NUMBER: 08/307,410
 FILING DATE: 16-SEP-1994
 APPLICATION NUMBER: 08/222,612
 FILING DATE: 16-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Fisher, Carlos A
 REGISTRATION NUMBER: 36,510
 REFERENCE/DOCKET NUMBER: GP94003.CP2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619-535-2807
 TELEFAX: 619-546-7929
 TELEX:
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 876 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULAR TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: internal
 ORIGINAL SOURCE:
 US-08-436-664-34

Query Match 3.2%; Score 128.5; DB 2; Length 876;
 Best Local Similarity 21.0%; Pred. No. 0.0053;
 Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;
 QY HFTLDGIRAFARVSKPQGFSDALISRRQILQAACETLSNPRRRY-NEGLDDE 157
 DB 107 HYAADLITGMARAERE--GFAVKVISCGRDLTQ--LASPVVETTKKGIIDIE 158
 QY 158 EATVITDV-----PMKVPG-----ALCVQEGGETIVLRVG 190
 DB 159 SYPEYVERKYLTPQIIVDLKGLMGKSDIPGVPGIGERTAKVLKQFTVENVASI 218
 QY 191 EALLKERLPKSPKQ--DVLVLMALAFDVSRA---MALPPDFITGYEFVEALKLQ 245
 DB 219 DEIKGEKTKENLQYRLAL-LSKQLAICRDPAVELTD--DIVYGEDEKRYVALFOE 275
 QY 246 EGASSLADPLAQIDR-----TLBETPRVYLELGL--PLGDDYAAKRING- 290
 DB 276 LGQSFIDKMAVQVDEGEKPLAGMDPAIADSVTDEMLADKAAVAVRVDGNYHAPIVGI 335
 QY 291 -----LSGVNIIISVGGGASALVNGLTREKPMNAFLMNTAAE----- 330
 DB 336 ALANERGFPLRPETALADPPLAM-----LQDDETKKTKMPDSKAAVALLKWKGI 385
 QY 331 -----QVLDLFAVATPSNIPAS-----FEYVEVALALVAQAFYIK-----KPHLQ 370
 DB 386 ELAGVVDLILAAVYLDPAQAAGVAVAVKHQVEAVRSDEAVVGKAKRTVPDEPTLAE 445
 QY 371 DADKQFOOLQAAKYAMEIPALYDTNNNE-IDFGIERGICALLIG-----KVDECRM 423
 DB 446 HLAH-----KAAAIWALEBPLMDELRRNQDRLLTELOPLAGILAMNEFTGVAVDTKRL 500

QY 424 -WGLD-SEDSQYRNPAIVEFVLENSNRDNDLFGCLKLETWLAGVPRFRDTRDKK 481
 DB 501 EQWAELETLQDQ-----AVERRIYELAQEFNINS--KQGT-----VLPDKQLPVLRK 549
 QY 482 PKGDIYDDPMVLSYL-ERVEVVOGSPLAAPATMARIGAEHVKASAMQLOKVPSPRTD 540
 DB 550 TKTG-YSTSDVLEKLPNHEIYVE--HILHYRQLGKQSTYIE-GLKVHAPVTKKVTM 605
 QY 541 RNSAEPRDVOET--VFSVDPVGNVNGDGEGRVIAEAVPSE-----NFEINDVAI 590
 DB 606 FNQA---LTQGLSSVPEPLQNIPIRLBERKIRQAFVSEPDMLIPADYSQIELRV 661
 QY 591 RAGVSESSVDETVE-----MSYADLKE--ASVKILAGAVALGILISFSQ 634
 DB 662 LAHIAE---DNLIEARRGIDITKTAMIDFHVSESDYANMRQAKAVNFGIVYGISD 718
 QY 635 KYFLKSSSPQRKDVSMESDVATIGSVRADSEALPRMDARPAENIVSKWQIKSLAF 694
 DB 719 -YGLAQMINITRKKAAPFIERFYFASPPGVK-----QYMDNIVOE-AKQKGYVT 764
 QY 695 GPDRIEMLEPVLDGRN-LKIMTDRA-ETAGLGVLVDTLLKLSVD-SVTVSADGTRA- 750
 DB 765 TLRHRRYLPDITSRNPNVSPARTAMTPIQSAD-ITKAMIDLSVRLREERLQAR 823
 QY 751 -----LYEATLESACLSDLVHPE 769
 DB 824 LLQVHDELILKAPKEIERLCRLV-PE 850

RESULT 9
 US-09-135-642-20
 : Sequence 20. Application US/09135642
 : Patent No. 6066483
 : GENERAL INFORMATION:
 : APPLICANT: RIGGS, MICHAEL G.
 : APPLICANT: SIVARAM, MATTHUR
 : APPLICANT: TUDOR, STARLA D.
 : TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
 : TITLE OF INVENTION: STEAROTHERMOPHILUS
 : NUMBER OF SEQUENCES: 34
 : CORRESPONDENCE ADDRESS:
 : ADDRESSER: Gen-Probe Incorporated
 : STREET: 9880 Campus Point Drive
 : CITY: San Diego
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 92121
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSeq Version 1.5
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/135,642
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/394,232
 : FILING DATE:
 : APPLICATION NUMBER: 08/307,410
 : FILING DATE: 16-SEP-1994
 : APPLICATION NUMBER: 08/222,612
 : FILING DATE: 16-SEP-1994
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Fisher, Carlos A
 : REGISTRATION NUMBER: 36,510
 : REFERENCE/DOCKET NUMBER: GP94003.CP2
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 619-535-2807
 : TELEFAX: 619-546-7929
 : TELEX:
 : INFORMATION FOR SEQ ID NO: 20:
 : SEQUENCE CHARACTERISTICS:

/ LENGTH: 876 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: internal
/ ORIGINAL SOURCE:
us-09-135-642-20

Query Match 3.24; Score 128.5; DB 3; Length 876;
Best Local Similarity 21.0%; Pred. No. 0.0053;
Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

```
QY 99 HFLTDGIRRAFEARVSKPPQFGSSDDALISRQILQAACTLSNPSRRRY-NEGILDE 157
DB 107 HYADDLIGTMAARARE--GFAVKVISGDRDLTQ---LASPVTVETTKKGLTDIE 158
QY 158 EATVITDV-----PMDKVPG-----ALCVLQEGGETEIVLRVG 190
DB 159 SYTPETVVEKYGTLTPEDIVDLKGLMGDKSNIPGVGIGKTAVKLLKQFGVENVLASI 218
QY 191 EALLKERLPKSPQ--DVIWMAIAFLDVSRA--MALDPDFITGYEFVEEALKULO 245
DB 219 DEIKGEKLEKENLQYRDIAL-LSKQLAAICRDAFVELTLD--DIVYKGEDEKRVALLFOE 275
QY 246 EGASSLAAPDLRAIDF-----TLSEITPRVYLELGL---PLGDDYAKRLNG- 290
DB 276 LGFQSFIDKMAVQTDGEEKPLAGMDFAIADSVTDEMADKAALVEVAGNYYHHAPIVGI 335
QY 291 -----LSGVNIIISVGGGASALVGLTREKPMNEAFLRMTAAE---- 330
DB 336 ALANERGFRLRPETALADKFLAW-----LQDERKKTKTMFSPKSAVALKTKGI 385
QY 331 -----QVDFVATPSPNIPAS-----FEVVEVALVAQAFIGK-----KPHLLQ 370
DB 386 ELRGVFDLLIAAYLLDPAQAAGVAVAIAKMGYEAVRSDAVGKAKRTVPDEPTLAE 445
QY 371 DADKQFOQLQAQVAMMEIPAMLYDRNNME-IDFGELGICALLIG-----KVDCRM 423
DB 446 HIAAR-----KAAIIMALEEPMLDELNRNEODRLTELEQPLAGILANMEFTGVAVTKRL 500
QY 424 -WLGLD-SEDSQYRNPAIVEFVLENSNRDNDLPGCLLETWLAGVFPFRFDTKDK 481
DB 501 EQWGAELTBLQ--AVERRIYELAGQFNINSP--KQLGT---VLFDKQLPVLKX 549
QY 482 FKLGDIYDDPMVLSYL-ERVEVVOGSPPLAAATMARI GAEHVAKASAMQALQKPPSRITD 540
DB 550 TKTG-YSTSADEVLEKLAIPHHEIVE--HILHYRQLGKLQSTYIE-GLIKVHPVVGKHYM 605
QY 541 RNSAEPDVOET--VFSVDVGNVNGRDSRGVFAVAPSE-----NPEINDYAI 590
DB 606 FNGA---LQGTGRLSVEEPNLQNIPIRLBEGKIRQAFVPSSEDDMLIPADYSQIELRV 661
QY 591 RAGVSSSVDETVE-----MSVADMLKE--ASYKILAGVAILGLISLFSQ 634
DB 662 LAHIAE---DDNIIEARRGIDITHTKAMDIPHVSEDDVANMRQAKANFGIVGISD 718
QY 635 KYLKSSSSQQRQKQWMSMESDVATIGSVRADSEALPRMDARTAEIVISKQKIKSLAF 694
DB 719 -YGLAONLINTRKAAEFIERYPASFGVK-----QYMDIIVDE-AKQKGYVT 764
QY 695 GPHRTIEMPEVLDGRM-LKIMTDRAA-ETRAQLGLVYDTLLKLSVQ-SVTVASDGRRA- 750
DB 765 TLHRRRRYLDITSRNENVSFAERTAMNPIQSSAAD-ITKAMIDLVSRLREBLQAR 823
QY 751 -----LVEATLIESACTSLDVHPE 769
DB 824 LLLQVHDELLEAPKKEIERLCRLV-PE 850
```

RESULT 10

US-09-135-642-32
/ Sequence 32, Application US/09135642
/ Patent No. 6065483
/ GENERAL INFORMATION:
/ APPLICANT: RIGGS, MICHAEL G.
/ APPLICANT: SIVARAM, MATTHEW
/ TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
/ TITLE OF INVENTION: STEAROTHERMOPHILUS
/ NUMBER OF SEQUENCES: 34
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Gen-Probe Incorporated
/ STREET: 9880 Campus Point Drive
/ CITY: San Diego
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 92121
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 1.5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/135,642
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/394,232
/ FILING DATE:
/ APPLICATION NUMBER: 08/307,410
/ FILING DATE: 16-SEP-1994
/ APPLICATION NUMBER: 08/222,612
/ FILING DATE: 16-SEP-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fisher, Carlos A.
/ REGISTRATION NUMBER: 36,510
/ REFERENCE/DOCKET NUMBER: GP94003.CP2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 619-535-2807
/ TELEFAX: 619-546-7929
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 32:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 876 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: internal
/ ORIGINAL SOURCE:
us-09-135-642-32

Query Match 3.24; Score 128.5; DB 3; Length 876;
Best Local Similarity 21.0%; Pred. No. 0.0053;
Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

```
QY 99 HFLTDGIRRAFEARVSKPPQFGSSDDALISRQILQAACTLSNPSRRRY-NEGILDE 157
DB 107 HYADDLIGTMAARARE--GFAVKVISGDRDLTQ---LASPVTVETTKKGLTDIE 158
QY 158 EATVITDV-----PMDKVPG-----ALCVLQEGGETEIVLRVG 190
DB 159 SYTPETVVEKYGTLTPEDIVDLKGLMGDKSNIPGVGIGKTAVKLLKQFGVENVLASI 218
QY 191 EALLKERLPKSPQ--DVIWMAIAFLDVSRA--MALDPDFITGYEFVEEALKULO 245
DB 219 DEIKGEKLEKENLQYRDIAL-LSKQLAAICRDAFVELTLD--DIVYKGEDEKRVALLFOE 275
QY 246 EGASSLAAPDLRAIDF-----TLSEITPRVYLELGL---PLGDDYAKRLNG- 290
DB 276 LGFQSFIDKMAVQTDGEEKPLAGMDFAIADSVTDEMADKAALVEVAGNYYHHAPIVGI 335
```

```

QY 291 -----LSGVRNIIIMSVGGGASALVGLTRKPKMNEAFLMTAA----- 330
DB 336 ALANERGFPLRPETALADPKFLAW-----LDDETKKTKMFKPSKRAAVALKMKGI 385
QY 331 -----QVDLFVATPSPNIPAES-----FEYEVALLVAQAFIGK-----KPHLLQ 370
DB 386 ELKGVAVDILLAAVLDLPQAAGVAVAAMKHQVEAVRSDBAVYGKAKRTVPDEPTLAE 445
QY 371 DADKQFOQLQAQKVMAMEIPAMLYDTNNWE-IDFGIERGICALLIG-----KVDECRM 423
DB 446 HLAAR-----KAAAIWALEBPLMDLRNEDRLTLEQPLAGILAMNEFTGVAVDTKRL 500
QY 424 -WGLD--SEDSQYRNPAIVEFVLNSNRDNDLPGCLKLETLAVGVPRPFDTDKX 481
DB 501 EOMGAELTEQLQ-----AVERRIYELAQEFNINSP-----KOLGT-----VLEDKQLPVLKX 549
QY 482 PKLDYDDPMVLSYL-ERVEVVGSPPLAAATWARIAGHVKASAMQALQKVPSPRYTD 540
DB 550 TKTG-YSTADVLEKLAIPHIEVE--HILHYRQGLQSTYIE-GLKVVHPVTKVHTM 605
QY 541 RNSAEPDVOET--VFSVDVGVNNVGDGEPGVFIABVPSF-----NFTNDYAI 590
DB 606 FNOA-----LTQTRLSSEBPNLQNIPIRLBGRKIROAFVPSBEDMLIFADYSQIELRV 661
QY 591 RAGVSESSVDETYE-----MSVADMLKE-----ASVYLLAGVAIGLISLSFQ 634
DB 662 LAHIAE---DNNLLEAFRRGLDHTTKTAMDIFHVESESDVTANNRQKAVNFGVYGISD 718
QY 635 KYFLKSSSFQKDMVSMESDVATIGSVRADSEALPRMDARTAENIVSKMOKIKSLAF 694
DB 719 -YELAQNMLNTRKEAEPFIERYPASFPQVK-----QYMDIIVGE-AKOKYVT 764
QY 695 GPDHRIEMLEPEVLDRM-LKIWTDRAA-ETAGLGLVYDTLLKLSVD-SYTVSADGTRA- 750
DB 765 TLHRRRYLPDITSRNFNVSFAERTAMTPIQSSAD-IIKKAMIDLSTVLRERLQAR 823
QY 751 -----LVEATLESACLSDLVHPE 769
DB 824 LLIQVHDELTEAPKEIERLCRLV-PE 850

RESULT 11
US-09-135-642-34
; Sequence 34, Application US/09135642
; Patent No. 6066483
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATHOOR
; APPLICANT: TUDOR, STARLA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,642
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/394,232
; FILING DATE:
; APPLICATION NUMBER: 08/307,410

```

```

; FILING DATE: 16-SEP-1994
; APPLICATION NUMBER: 08/222,612
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fisher, Carlos A
; REGISTRATION NUMBER: 36,510
; REFERENCE/DOCKET NUMBER: GP94003.CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-2807
; TELEFAX: 619-546-7929
;
; TELER:
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 876 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-09-135-642-34
;
; Query Match
; Best Local Similarity 21.0%; Pred. No. 0.0053;
; Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;
;
; 3.2%; Score 128.5; DB 3; Length 876;
;
; 99 HPLTDGIRRAPEARVSKPPQFGSDMLISRRQLQAACETLSNPSRRRY-NEGGLDDE 157
; 107 HEADLDITGMAARAERE--GFAVKVISGDRDLTQ-----LASPVTEITKKGIIDIE 158
; 158 EATYITDV-----PMDKVPG-----ALCYOEGSETEIVRVG 190
; 159 SYRPEIVVEKYGTLTPEQIVDLKGLMGKSDNIRVPEIGETKAVKLGQTEYENVLASI 218
; 191 EALKERLPKSPFO--DVVLWMLAFIDVSRDA---MALDPPFITGEFVEEALKLLQE 245
; 219 DEIKGELKENLQRYRLAL-LSKQLAICRDAPVELTLD-DIVYGEDEBKVALFOE 275
; 246 EGASSIAPDLRAQIDE-----TLBEITPRVYLELGL--PLGDDYAAKRLNG- 290
; 276 LGFQSFIDKMAVQTDSEBEKPLAGMDFAIADSVTDEMADRAALVVEVVGDNVHNAPIVGI 335
; 291 -----LSGVRNIIIMSVGGGASALVGLTRKPKMNEAFLMTAA----- 330
; DB 336 ALANERGFPLRPETALADPKFLAW-----LDDETKKTKMFKPSKRAAVALKMKGI 385
; QY 331 -----QVDLFVATPSPNIPAES-----FEYEVALLVAQAFIGK-----KPHLLQ 370
; DB 386 ELKGVAVDILLAAVLDLPQAAGVAVAAMKHQVEAVRSDBAVYGKAKRTVPDEPTLAE 445
; QY 371 DADKQFOQLQAQKVMAMEIPAMLYDTNNWE-IDFGIERGICALLIG-----KVDECRM 423
; DB 446 HLAAR-----KAAAIWALEBPLMDLRNEDRLTLEQPLAGILAMNEFTGVAVDTKRL 500
; QY 424 -WGLD--SEDSQYRNPAIVEFVLNSNRDNDLPGCLKLETLAVGVPRPFDTDKX 481
; DB 501 EOMGAELTEQLQ-----AVERRIYELAQEFNINSP-----KOLGT-----VLEDKQLPVLKX 549
; QY 482 PKLDYDDPMVLSYL-ERVEVVGSPPLAAATWARIAGHVKASAMQALQKVPSPRYTD 540
; DB 550 TKTG-YSTADVLEKLAIPHIEVE--HILHYRQGLQSTYIE-GLKVVHPVTKVHTM 605
; QY 541 RNSAEPDVOET--VFSVDVGVNNVGDGEPGVFIABVPSF-----NFTNDYAI 590
; DB 606 FNOA-----LTQTRLSSEBPNLQNIPIRLBGRKIROAFVPSBEDMLIFADYSQIELRV 661
; QY 591 RAGVSESSVDETYE-----MSVADMLKE-----ASVYLLAGVAIGLISLSFQ 634
; DB 662 LAHIAE---DNNLLEAFRRGLDHTTKTAMDIFHVESESDVTANNRQKAVNFGVYGISD 718
; QY 635 KYFLKSSSFQKDMVSMESDVATIGSVRADSEALPRMDARTAENIVSKMOKIKSLAF 694

```

Db 719 -YGLAQLNLTTRKEAAEFIRYFASFGVX-----QYMDNIYOE-AKQKGYVT 764
Qy 695 GPDHRIEMLEPEVLDDGRM-LKIWTDRAA-ETPAQLGLVYDYTLTKSVQ-SVTVASDGTBA- 750
Db 765 TLHRRRYLPDITSRNPNVNSFAERTAMTPPIQSAAD-IKKAMIDLSVRLREERLOAR 823
Qy 751 -----LVEATLEBSACTSLDYHPE 769
Db 824 LLLQVHDELILEAPKEIERLCRLV-PE 850
RESULT 12
US-08-394-232A-20
Sequence 20, Application US/08394232A
Patent No. 6100078
GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: SIVARAM, MATHOOR
APPLICANT: TUDOR, STARLA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHERMOPHILU
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gen-Probe Incorporated
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,232A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 876 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHECAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-394-232A-20
Query Match 3.2%; Score 128.5; DB 3; Length 876;
Best Local Similarity 21.0%; Pred. No. 0.0053;
Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;
Qy 99 HPLTDGIRAFARVSKPQFGSFDALISROTLQAACETLSNPRSRREY-NEGILDD 157
Db 107 HYAADLIGTMARARE--GPAVKYISGRDLITQ-----LASPYVTEITTKGIDIR 158
Qy 158 EATVITDV-----PMDKVPG-----ALCVLQEGGETEIVLRVG 190

Db 159 SYTPEYVERKYGLTPBOIVDLKGLMDKSDNIGVPGIGERTYVKKLKQGTVENYVAST 218
Qy 191 EALIKERLPKSPKO--DVVLWMLAFLDVSRD---MALDPDFITGVEFEVALKILQ 245
Db 219 DEIKGEKLENLQRYDIAL-LSKQLAICRDAPEVLTLD--DIVYKGEDEKRYVALFOE 275
Qy 246 EGASSIAPDLRAQIDE-----TLEETPRYVLEILGI---PLGDDYAAKRLNG- 290
Db 276 LGFQSPFLDKMAVQTDSEKPLAGMDPALADSVYDEMILADKAAVLVEVGDYHNAPIVGI 335
Qy 291 -----LSGVRNILNSVGGGAGALVGLTRKFMNEAFIRMTALE----- 330
Db 336 ALANERGFPLRPETALADPFLAW-----LGDEYKKTTFDSKRAVALKMKGI 385
Qy 331 -----QVDFVATPSPNIPAS-----FEYEVALLVAQAFIGK-----KPHLQ 370
Db 386 ELRGVVPDLILAAYLDPADQADVAAVAKMQYEAVRSDEAVYGAKAKRTVDEPTLAE 445
Qy 371 DADKQFQQLQAAKVMWEIPAMLYDRNWE-IDFGLERGLCALLIG-----KVDECRM 423
Db 446 HLAR-----KAAIWMLEBPLMDLRNEDRLITLEQPLAGILANMEFTGVVDTKRL 500
Qy 424 -WIGLD-SEDSQRYNPAIVEFVLENSNRDNDLPLGLCKLETMTWLAGVPRPRDTDKX 481
Db 501 EOMGAEITBQLQ---AVERRIYELAQEFNINSP---KQLGT-----VLPDKLQLPVLRK 549
Qy 482 FKLDYDDPMTLSYL-ERVEVYQSSPLAAATMAIRIGAEHVASAMQALQKVPSPRYTD 540
Db 550 TKYG-YTSDADVLEKLAHPHIEVE--HILHYRLGLQSTYIE-GLLKVHPPTGKHTM 605
Qy 541 RNSAEPDVQET--VPSVDVGNNGRDEPGVFIABAVPSE-----NEFTNDYAI 590
Db 606 FMOA---LTQTRLSSVEBNLQNIPLRLEEGAKIRQAFVPSBDWLIFADYSQIELRV 661
Qy 591 RAGVSSSVDETVE-----MSVADMLKE--ASVKILAGALIGLISLSFQ 634
Db 662 LAHIAE---DDNILEAFRGLDITHTAMDIFHVSEBDVTANRRQAKANFQIVYGISD 718
Qy 635 KYFLKSSSFQRDMVSMESDVATIGSVRADSEALPRMDARTAEIVSKMKISLAF 694
Db 719 -YGLAQLNLTTRKEAAEFIRYFASFGVX-----QYMDNIYOE-AKQKGYVT 764
Qy 695 GPDHRIEMLEPEVLDDGRM-LKIWTDRAA-ETPAQLGLVYDYTLTKSVQ-SVTVASDGTBA- 750
Db 765 TLHRRRYLPDITSRNPNVNSFAERTAMTPPIQSAAD-IKKAMIDLSVRLREERLOAR 823
Qy 751 -----LVEATLEBSACTSLDYHPE 769
Db 824 LLLQVHDELILEAPKEIERLCRLV-PE 850
RESULT 13
US-08-394-232A-32
Sequence 32, Application US/08394232A
Patent No. 6100078
GENERAL INFORMATION:
APPLICANT: RIGGS, MICHAEL G.
APPLICANT: SIVARAM, MATHOOR
APPLICANT: TUDOR, STARLA D.
TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHERMOPHILU
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gen-Probe Incorporated
STREET: 9880 Campus Point Drive
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS


```

SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,232A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/307,410
FILING DATE: 16-SEP-1994
APPLICATION NUMBER: 08/222,612
FILING DATE: 16-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fisher, Carlos A
REGISTRATION NUMBER: 36,510
REFERENCE/DOCKET NUMBER: GP94003.CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-2807
TELEFAX: 619-546-7929
TELEX:
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 876 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-394-232A-32

```

Query Match 3.2%; Score 128.5; DB 3; Length 876;

Best Local Similarity 21.0%; Pred. No. 0.0053;

Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

```

99 HFLTDGIRRAFEARVSKPQGFSDDALISRQILQACETLSNPSRREY-NEGLLDDE 157
107 HYBADDIIGMAAABRE--GFAVKVLSGRDLTQ-----LASPOVTEITKGIIDIE 158
158 EATYITIV-----PMKVPG-----ALCVLOEGGETIVLRVG 190
159 SYPTETVEKYGTLPEQIVDLKGLMGKSDNIPGVIGEKTAVKLKQFETVENVASI 218
191 EALKERLPKSPKQ--QVVLVMAALFIDVSRD--MALDPPDITGVEFEALKLOE 245
219 DEIKGEYTKERLRYDLAL--LSKOLAICRDAPVELTLD--DIVYGEDEKVALFQE 275
246 EGASSLAPDLRAQIDE-----TLERITPRVYLELGL--PLGSDYAKRLNG- 290
276 LGPQSTLDKRAVQTDDEBEKPLAGMDFAIADSVTDEMADKALVYEVVGNHYHAFLVGI 335
291 -----LSGVNNIISVGGGASALVGGLTREKEMNEAFLMTAAE--- 330
336 ALANERGFPLRPETALADPKFLAM-----LGEETKKTMPDSKRAVALKWKGI 385
331 -----QVDLVATTSNTPAS-----FEYVYVALLVQAQFIGK-----KPHLLQ 370
386 ELRGVVFLLLAAYLDPQAAGVAAVAKHQYEAVRSDAVYGAKATVPDEPTLAE 445
371 DADKQFOLOQAQKVMAMEIPAMLYDTNNME-IDFGELGICALLIG-----FYDECRM 423
446 HLAAR-----KAAAIWALEEPLMDLRNBOURLTELEQPLAGILAMFEFGVAVDTKRL 500
424 -WIGLD-SEDSQYRNPAIVEFVLNSNRDNDLPGLCLELTYLAVGVPFRFDTQDK 481
501 EOMGABELEQIQ-----AVERIYELAGQEFINSP-----KQIGT-----VLFKQLPLVKK 549
482 FKLDGYDDDMVLSYL--ERVEVVGSGPLAAATMARIKAHEVKSAMQALQKVPSPRYTD 540
550 TKTG-YSTGADVLEKLAIPHHEIVE--HILAYRQKGLQSTYIE--GLIKVAVPYGVKVTM 605
541 RNSAEPRDVOBT--VFSVDVPVGNVGRDGEKGVPIAAVPRSE-----NFEETDYAI 590

```

```

606 ENQA---LITQGLSSVEBNLQNIPIRLBGRKIROAFVPSPEPWLIIPADYSQIELRV 661
591 RAGVSESSVDETTYE-----MSVADWLKE---ASVKILAAAGVAILGISFSQ 634
662 LAHIAE---DNLLEARRRGIDITHTKAMDIHFVSEEDVATNMROKAKAVNFGVYGISD 718
635 KYFLKSSSFQKQKQVMSMESDVATIGSVRADSEALPRMDARTAEINIVSKQKIKSLAF 694
719 -YGLAQNLNITRKAAPAFIERFYAFSPGVK-----QYMDNIVQE-AKQKGYVT 764
695 GPDHRIMLPEVLDDGRN-LKIWTDRRA-ETROGLVYDYLTKLSVD-STVSGADGIRA- 750
765 TLHRRRYLPDITSRNPNRSFARTMTPIQSSAD-IKKAMIDLSTVRLREERLQAR 823
751 -----LVEATLEESACISDLVHPE 769
824 LLLQVHDELILEAPKEIERLCRLV-DE 850

```

RESULT 14

US-08-394-232A-34

; Sequence 34; Application US/08394232A

; Patent No. 610078

; GENERAL INFORMATION:

; APPLICANT: RIGGS, MICHAEL G.

; APPLICANT: SIVARAM, MATHOOR

; APPLICANT: TUDOR, STARLA D.

; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHERMOPHILUS

; NUMBER OF SEQUENCES: 34

; CORRESPONDENCE ADDRESS:

ADDRESS: Gen-Probe Incorporated

STREET: 9880 Campus Point Drive

CITY: San Diego

STATE: CA

COUNTRY: USA

ZIP: 92121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/394,232A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/307,410

FILING DATE: 16-SEP-1994

APPLICATION NUMBER: 08/222,612

FILING DATE: 16-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Fisher, Carlos A

REGISTRATION NUMBER: 36,510

REFERENCE/DOCKET NUMBER: GP94003.CP2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-535-2807

TELEFAX: 619-546-7929

TELEX:

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:

LENGTH: 876 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

US-08-394-232A-34

Query Match 3.2%; Score 128.5; DB 3; Length 876;

Best Local Similarity 21.0%; Pred. No. 0.0053;

Fri Jun 10 09:57:54 2005

us-10-600-070-2.ra1

Page 13

Job time : 33 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2005, 00:53:22 ; Search time 26 Seconds
(without alignments)
2964.216 Million cell updates/sec

Title: US-10-600-070-2
Perfect score: 4063
Sequence: 1 MEALSHVIGLSLSPFQICRLP.....YEVFWKSGWKTEGSVLAS 801

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	497	12.2	798	2 AD2144	hypothetical prote
2	429	10.6	714	2 S7082	hypothetical prote
3	145	3.6	2297	2 AB2494	hypothetical prote
4	143	3.5	2089	1 A48757	acetyl-CoA carboxy
5	139.5	3.4	879	2 JC4286	DNA-directed DNA P
6	138	3.4	2021	2 A84771	similar to ch-TOG
7	131.5	3.2	1194	1 DUB28	DNA-directed DNA P
8	131.5	3.2	2111	2 A7068	mycocerosate synth
9	130.5	3.2	1026	2 C97783	cell surface antig
10	129.5	3.2	2110	2 B44110	mycocerosate synth
11	129	3.2	1275	2 T49362	hypothetical prote
12	128.5	3.2	836	2 T30312	pylin biosynthetic
13	128.5	3.2	876	2 S70368	DNA polymerase I -
14	127.5	3.1	2472	2 E83594	acetyl fructose 1 -
15	126.5	3.1	1018	2 T30853	antigenic heat-sha
16	125.5	3.1	718	2 A81122	hypothetical prote
17	125.5	3.1	4385	2 T29042	hypothetical prote
18	123.5	3.0	673	2 C63080	probable chemotaxi
19	123.5	3.0	848	2 C70834	probable endopepti
20	123	3.0	3421	1 W2886	36k tegument prot
21	122.5	3.0	899	2 H87513	hypothetical prote
22	122.5	3.0	2756	2 T30183	hypothetical prote
23	121	3.0	1381	2 AF2010	regulatory protein
24	120	3.0	1252	2 D72122	RNA polymerase bet
25	120	3.0	1262	2 F81548	DNA-directed RNA P
26	120	3.0	3938	2 T42761	Bassoon protein -
27	119.5	2.9	692	2 G87415	glycyl-tRNA synth
28	119.5	2.9	698	2 A82593	hypothetical prote
29	119.5	2.9	831	2 S26675	DNA-directed DNA P

30	118.5	2.9	3942	2 T42730	Bassoon protein -
31	117.5	2.9	1004	2 G87323	hypothetical prote
32	117.5	2.9	4613	2 T17409	polyketide synthas
33	116.5	2.9	880	2 B69680	DNA polymerase I p
34	116.5	2.9	1074	2 T01884	hypothetical prote
35	116.5	2.9	1644	2 AC0823	probable lipoprote
36	116.5	2.9	3450	2 T26963	hypothetical prote
37	116.5	2.9	3461	2 T26964	hypothetical prote
38	115.5	2.8	693	2 G82618	pluss biogenesis p
39	115.5	2.8	804	2 T44506	transducer protein
40	115	2.8	1983	2 AC1922	two-component hybr
41	114.5	2.8	758	2 A98290	exopolysaccharide
42	114.5	2.8	758	2 A98290	hypothetical prote
43	114.5	2.8	1755	2 F82618	chemotaxis-related
44	114.5	2.8	2333	1 GNNY2F	genome polyploid
45	114	2.8	777	2 T44597	transducer protein

ALIGNMENTS

RESULT 1

AD2144 hypothetical protein al12707 [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. strain PCC 7120

A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C/Accession: AD2144

R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasaoka, S.; Watanabe, A.; Iriuchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, M.; DNA Res. 8, 205-213, 2001

A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat

A/Reference number: AB1807; MUID:21595285; PMID:11759840

A/Accession: AD2144

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-798 <KUR>

A/Cross-references: UNIPROT:Q8YTL0; GB:BA000019; PIDN:BA074406.1; PID:G17131800; GSPDB:G

A/Experimental source: strain PCC 7120

C/Genetics:

A/Gene: al12707

C/Superfamily: Nostoc sp. hypothetical protein al12707; dnaJ amino-terminal homology

Query Match	12.2%	Score 497;	DB 2;	Length 798;
Best Local Similarity	24.6%	Pred. No. 6e-24;		
Matches	207;	Conservative 139;	Mismatches 301;	Indels 196; Gaps 33;
QY	84	VPIIDFYOVVGAQTHPLTDGIRRAFEARVSKPQFGSPDALTISRQIIQACETLSNP	143	
DB	11	VRIPLDYRILGLPLASDQLRQAYSDRIYQLPFRREYSQALISRRQLEBAVVLSDP	70	
QY	144	RSRRRY-----	NEGLDDEAVITVDVPMKYPGALCV	176
DB	71	KERSSYQOLYLAAHYDDNATTKVAVENRGDSNNGHFDVQSLSI--EVSSEELIGALLI	128	
QY	177	LOEGGTEIVLRVGEALL-----	KERLPKSF-----KQDVVLWALAFIDVSR	219
DB	129	LOEGEYELVVKQGRNLTGNQNGTASTRNGNHRPESFLDSSEPDLLTVALASLEIGR	188	
QY	220	D-----	AMALDPPDFITGYEFVEBALKTLQEBGASLAPDLRAQIDETLEETPR	269
DB	189	EQWQGHYENALSLF-----	TQGE-----VLFSGS-----IPFSVQAEIQADLYKLRPY	233
QY	270	VYIELGLPLGDDVAAKRLGLSGVRNIIISVGS--GASALVGLTBE--KFMNEAFR	325	
DB	234	RIIEELALP--QEKTIERRHQLDLSIILDDRGIDDTGNDOSGLNIDDFLRTOQRHN	291	
QY	326	MTAAEQVDLFAVAPSNIPASFEYVEVALVQAIFGKKPHLLQADKQOQOQAKVM	385	
DB	292	LTVAEQKLPDG--ESKSPS--AVATYTLAVVASINAGFQROKALTRNAKQILMRLSKO--	347	
QY	386	AMEIPALVDYTRNNWEIDFGLEGLCALLIGKVDCEHMTGLDSEDSQYRNPAIVEPVL	445	

```

Db      348 -----DVHLEQSLCALLLGQTEBEATRVLELSQE---YEALALI---R 383
Qy      446 NSNRDNDNDLPGICKLETWLAGVFPFRDTPKDKFKLGDYYDDPWLSTYLERV----- 500
Db      384 EKSQDSBDLLPGICLVAEQVLQNEVFPHPFDLSQGLSKDYRANQOVQVLEALPDAE 443
Qy      501 -----EYVQGSPLAAATMAR--IGAEN--VKASAMQLOKVPSP---RYTD 540
Db      444 TTMEVAVINRQSPQRPGRNSYSGTVPAAKRPVGKANRPGASTPVPQRSHSPSVNQFH 503
Qy      541 RNAAEKDQVQETVPSPDVPVNNVGRDSE-----PGVFIAEAVRPSNFEPTND 587
Db      504 QNTTPELEPTNSNHRPSSNFTTARENISTTDAYTDNYPPELPVERASRP----- 555
Qy      588 YAIRAGVS---ESSVDETTV-----EMSVADMKEASVKILAA 622
Db      556 --VQPVSGTQSTPRPQTKRRRRKKRQAVNVRGSHIQOQRPSTLGRKRLMIVL 613
Qy      623 GVAIGLISLF---SOKY-FLKS---SSSFQRKDVVSMESDVATIGSVRADSEALPR 673
Db      614 G-SLGGILFMLIVSTTFGWLKXVFPFAPSLQGEQLSIQSPPLEIPDKNAQIQSPES 672
Qy      674 MDARTAEIVSKKQKIKSLAFGPDHRIEMLPVULDGMKLTIDRAETQQLGVYDTYL 723
Db      673 LTETARKIIEMLATYASALGAEHKIESINEILTGSALQWMLIALQDRAENRHRYS- 721
Qy      734 IKLVSQSVTVS-ADGTRALVEATLEESACSLDIYH--PENNAVDVRYTYRVEVFMKSG 790
Db      732 HSKVDISIGSDIDPRASGATVRE---LTQPEYENQKKSGBER-LRRRYELIRDDI 787
Qy      791 WKT 793
Db      788 WRI 790

RESULT 2
S76082
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S76082
R:Kaneko, T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamiu, E.; Nakamura, Y.; Miyajima, N.;
O., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76082
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-714 <KAN>
A:Cross-references: UNIPROT:Q55553; EMBL:D63999; GB:AB001339; NID:g1001396; PIDD:BA11006
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Start codon: GTG
C:Superfamily: Nostoc sp. hypothetical protein a112707; dnaI amino-terminal homology
F:6-70/Domain: dnaI amino-terminal homology <DMU>

Query Match      10.6%; Score 429, DB 2; Length 714;
Best Local Similarity 25.1%; Pred. No. 1.2e-19;
Matches 194; Conservative 124; Mismatches 331; Indels 124; Gaps 30;

```

```

Qy      86 IPIDFYOVLAGQTHFLTDGIRBAFEARVSKPQFGSDDLISROILDAACETLSNPRS 145
Db      3 IPIDFYRIIGTIPROSGGETTEQAYODRLQLPRKESDAATVTLRNQMLDAYETLRDPEK 62
Qy      146 RREYNE---GLTDE--EATVITDVPMDKVP---GALCVLOSGGETEYLVARGEALLKE 196
Db      63 RQAVDDQEWGAMDBALGEALPLTPPELESGPDEGTGALLTLIDGVELVAVKKGEPALHD 122
Qy      197 RLPKS--FKQDVVLVMAALFLDVSRDAMALDPDPFTIGYFVEEA-LKLQEGSGASTLAP 253

```

```

Db      123 PNPPAGGLPDVILLSTVLAHWEISRRWQOQ-----YEPATASIKALARLQOQNDPF 176
Qy      254 DLRAQIDETLEETIPRYVELLGLP-LGDDYAAKRLNGISGVNIIIMSVGG-GGASALVG 311
Db      177 ALBAEIRQELRYRPRRIEELLAKEGOGEE---QROGGLALLQAMQDRGIGIGKGDYS 223
Qy      312 GLTRE---KFMNAFLMTAAEQVDLFPATPSNIPABSEFVEYVALVAQAETGKKPHL 368
Db      234 GLQNDPFLKFIHQRLCHLTVAEQNALFL--PESQPSLVASYLAIVSHLMAGVEKQDPMA 291
Qy      369 LQDADPQFQLOQAAXMAMEIRPMLVDTNNWIDGLEGLGALLIGKVDCEGMMLGLD 428
Db      292 IYBAKKLITIQLENCQ-----DLAEKVITCELLHGGQTEV--VLAALD 330
Qy      429 SEDSQYRNPAIVFVENSNRDNDLPGICKLETWLAGVFPFRDTPKDKFKLGDYX 488
Db      331 QGD-----PKIVA-GLESKATGEDPLTAFTTEQMLEEIEIPYRDLSPFLSPRAYF 384
Qy      489 DDPWLSTYLERVEVQ-----GSP-LAAATMABIGAEHVKASAM-QALQKVPSPRYT 539
Db      385 NNPSVOQYLEQLBPDSTFTDINSFASPALSTATSETPMVHSGAALPDRRLTSTVRS--- 441
Qy      540 DRNSAEPKQVQETVPSPDVPVNNVGRDGERGVIAEAVRPSNFEPTNDVIRA----- 592
Db      442 -RRGRSPRSRDVFP-----SADNSGLAVT-TLSALAYDTHSLGTNGIGGDSTIS 491
Qy      593 -GVSESSVDETTYE-MSVADMKEASVK-----ILAGVAILGLISLFSQK----- 635
Db      492 NGFSSNSABESTSKHSPRRKKRVITIKPRPGIFLCLAGIYGAATALLINNTGDPILG 551
Qy      636 -----YFLKSSSFQRKDVVSMESDVATIGSVRADSEALPRDARTAEIVSKKQK 688
Db      552 LLEDPLDVFL-----DQPEFTPDEATSNLILISQ---PNNQOQGVQVVOGWLID 598
Qy      689 IKSLARGPDRRIEMLPVLDGRMLKITWDRAAETAGLVYDVTLLKLSVDSVTVS-ADG 747
Db      599 SKGLAFQGYNDGALGSVLAIPVLAQQRK-AQRDAQKVHYQHEKQITLAIQVNPDP 657
Qy      748 TRALVEATLEESACSLDIHPENNAVDV-RYTYRVEVFMKSGMKITGESVYL 799
Db      658 NRATVYARVEISQPTLGNQOKGSATDDLVVRQVLRHGQVWKIDQIOVV 710

RESULT 3
AB2494
hypothetical protein a117130 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a1y
C:Species: Nostoc sp. PCC 7120
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AB2494
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sawamoto, S.; Watanabe, A.; Iriyuchi,
N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Ana
A:Reference number: AB1807; MUID:2159285; PMID:11759840
A:Accession: AB2494
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2297 <KIR>
A:Cross-references: UNIPROT:Q8YU08; GB:BA000020; PIDD:BA878214.1; PID:g17135668; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: a117130
A:Genome: plasmid

Query Match      3.6%; Score 145, DB 2; Length 2297;
Best Local Similarity 19.6%; Pred. No. 1.5; 273; Indels 170; Gaps 29;
Matches 136; Conservative 114; Mismatches 147;

```

```

Qy      89 DRYOVLAGQTHFLTDGIRBAFEARVSKPQFGSDDLISROILDAACETLSNPRSRE 148
Db      273 ERYQIGLAT-----IAHVNEP-----RRREILKKALDVVAKTGSTG 311

```



```

Qy 149 YNE-----GLDDEEATVITDVPMDKVPGLCVLQEGGETEIVLGVGALLKER 197
Db 312 FISEVENTISLNLALDLKDBERTFVPOALBALPEKI-----ENGVN--IFETIAPQLTER 365
Qy 198 LPSFKODVVLVMAALFLVSRDAMALDP-----DPTGYEVEEALKILOEE-----G 247
Db 366 EPOLKEKTEIVNNLERGSDYQREALVAIAPLPKSEPIILQOALEIAVKIANEYDRIKA 425
Qy 248 ASSIAPDLRAQ-----IDETLEETPRVVELLGLPLGDDVAARINLSCVNRILW 299
Db 426 LAALAPLKKETQCNELVYLILERIENIWPNI-----SGVSETLV 465
Qy 300 SVGGGASALVGLTRREKFMNEAFLRMTAAEQVDLFVATPSNIPAE--SFEYVEVALALVA 358
Db 466 FI-----ASYL--SKSGSKLQKAF-----NIYQNEFGSYDADILVA 502
Qy 359 QA-FIGK-KPHLLQDADKQFOOLQO--AKVMA-----EIPAMLYDTRNMEI 402
Db 503 LAPHLSTLEBPQLQOALKIAINIKEGASKVVALPAVIPHLPQSKQLLEKAFETIQTLEY 562
Qy 403 DFGLEKGLCAL-----LIGKTVDECMWLGIDSEDSQYR--NPAI 439
Db 563 DNARSRLVAVIPIHLSDFESHLDKSLERIKLVSTYDSAKA-QALVAVAPQLKKFNPSTL 621
Qy 440 VEFVLE-----NSNRDNDLPGLC-----KLLETWLAGVFPFRPTDKDKKFKLGD 486
Db 622 LQALALEIAKKINSDEQDEALMAIAYOLSESEPELLBOYKIAQNTYTHSHKRIALIP 681
Qy 487 YDDPMLVSLERVEVVOGSPPLAAATMARIAGEHVAKASAMQ-ALQKVPSPRYTDRNSAE 545
Db 682 YLRBPQGTETFLQKFSATASIASAPCL--PEPKSEVLYQALKMLLDANYSDILIRAN 738
Qy 546 PKVOGEVFSVDPRGNVNGDGERGVFLAEAVRSENPETMDVAIRAGVSSSDVETTV- 604
Db 739 ELKTIYVLELKT-----EQKEVIEIAKTIK-NDTFKGAELIAFAVATHLSEPDQYKIF 790
Qy 605 EMSVADMLK--EASVKILAGVAIGLISLPSQKYPFKSSSFQKQMVSMESDVAITIG 661
Db 791 FLEITEHLKAIENDSSQAKALAVIPIYLSKSSPESLDBKAFIAENLQYQSCFPDDLVTILA 850
Qy 662 S-VRADSEALPRMDARTAEINIVSKQIKSLA 693
Db 851 THLKERECTKLLEQALKAKADIDSEYQAOQDFA 883

```

RESULT 4

A48757 acetyl-CoA carboxylase (EC 6.4.1.2) - Cyclorella cryptica

C/Species: Cyclorella cryptica
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: A48757

R/Roeseler, P.G.; Ohlrogge, J.B.

J. Biol. Chem. 268, 19254-19259, 1993

A/Title: Cloning and characterization of the gene that encodes acetyl-coenzyme A carboxy

A/Reference number: A48757; MUID:93374903; PMID:8103514

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-2089 <ROE>

A/Cross-references: UNIPROT:Q3478; GB:L20784; NID:G1065903; PIDN:AAA81471.1; PID:G40945

A/Note: authors translated the codon GGC for residue 1834 as Ala

C/Genetic: 25/1; 729/1

A/Intctions: 25/1; 729/1

C/Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-

C/Keywords: biotin binding; lipase

P/99-603/Domain: biotin carboxylase homology <BCH>

P/731-603/Domain: lipoyl/biotin-binding homology <LPB>

P/770/Binding site: biotin (lys) (covalent) #status predicted

Query Match 3.5%; Score 143; DB 1; Length 2089;
 Best Local Similarity 20.0%; Pred. No. 1.8;
 Matches 169; Conservative 113; Mismatches 289; Indels 272; Gaps 44;

```

Qy 73 LPSIDPRBHVPIPIDFYQVLAQTHLTDGIRRAFEARVSKPP-----QGFSDAL 126
Db 197 LPNALDK-----LGIFIGPTGVMSVLGDKIANIILAQAKVPSIPMSGSFGGPDGP 250
Qy 127 ISRRQIIQAACETLSNPRSREYNEGLD--DEATYITDVPMDKVPGLCVLQEG-- 181
Db 251 LQ-----ADLTEEGTIPMEIFPKGLVTSADAEVIAANKGWEN--GIMTKASEGGGCK 301
Qy 182 -----ETEIYLR-----VGEALLKERLPSFQDVTVLVMAALFLVSRDAMALDP 226
Db 302 GIRFVNDADIRNAFVOVSNELTSPFLMQLCKNAHIEVQIVG---DQGNAAVING 357
Qy 227 PDFTGYEVEVEALKILOEBGASLADPLRAQIDETLEETPRVVELLGLPLGDDVAAK 286
Db 358 RDCSTQRRF-----QKIFEEGSPSIVP-----KETHEM-----ELAAQ 391
Qy 287 RLNLGSGVRNIIWSVGGGASALVGLTRREKFMNEAFLRMTAAEQVDLFVA--TPSNIP 343
Db 392 RL-----TONIGYQ--GAGTVEYLYNADNKF--FLELNPRLQVHEPVTGEGITGANLP 441
Qy 344 ABS-----FEVEVALALVAQAFIGKKPHLLQDADKQFOOLQOAKVMAEIPA-- 391
Db 442 ATQLOVAMGIPLFNIPDIRLYGREDAYGDP--IDFLQEKYREL-DSHVIARITLENP 498
Qy 392 -----MLYDTRNM--EIDFGLEKGLCALLIGKVCRCMWLGIDSEDSQY-- 434
Db 499 DEGFKPTSGSIERIKFQSTPRVWGYFSVGANGI-----HEFPADSQGH 542
Qy 435 --BNP-----AIVEVLENSNRD--NDDLPGCLKLET-----WLAGVFP 472
Db 543 LFAKGPAREQARKALVIALKEMEVRGDIRNSVEYLVLLTEAEKQNTIDTSMIDGIT-- 600
Qy 473 RFRRTKQKPKRLGDYDDPMLVSLERVEVVOGSPPLAAATMARIAGEHVAKASAMQ-- 529
Db 601 ----KEKSYKV-----EMPSHL-----VVGAIVKFAFEHVAVATEEVKESFRKGVS 644
Qy 530 -----LQKVPSPRYTDRNSAEKDVQET----- 552
Db 645 TAGIPGINSRNIEBAVYIDTKPRIVERISPDVYFTLDGNTIDVEVQYTAGALLATFGG 704
Qy 553 ----VFSVD--PVGNNGVRDGE---PGVFIAEAVRPS-----ENFETNDVAIRAGVSES 597
Db 705 ETRHIFGMDPLGLRLSLDGAIVLMPITFDPSRLRTDVTGKVAVYLDNGATVBAQ---- 760
Qy 598 SVDETTEMSVADMLKASVKILAGVAIGLISLPSQKYPFKSSSFQKQMVSMESD SD 656
Db 761 ----QPYVEVAMKMI--MPIKATESG-----KITHNISAGSVISAGDLLASLEIKD 806
Qy 657 VATIGSV-----RADSEALPRMDARTAEINIVSKQIKSLAGPDRHIEMLPEVLDR 710
Db 807 PSRYKAIETPSGKLDINESKVDLEPQKAVNANVL-----GFNLD-----PEAVAAQ 852
Qy 711 MLKIWTDRAEIQAQGLVYD-YTLKLKLSVDSVTVSADGTRALVEATEESACSLDVHPE 769
Db 853 AIDBATDSAAADLVQVLEDFYAVESQPDGV-IADVVVRLITRANNE-----TLDDVISE 907
Qy 770 NNA 772
Db 908 NLA 910

```

RESULT 5

JC4286 DNA-directed DNA polymerase (EC 2.7.7.7) - Bacillus stearothermophilus

N/Alternate names: DNA polymerase I

C/Species: Bacillus stearothermophilus

C/Date: 14-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 17-Mar-1999

C/Accession: JC4286

R/Phang, S.M.; Teo, C.Y.; Lo, E.; Wong, V.W.T.

Gene 163, 65-68, 1995

A/Title: Cloning and complete sequence of the DNA polymerase-encoding gene (BspPolI) and

A/Reference number: JC4286; MUID:96001245; PMID:7557480

A/Accession: JC4286

Db 906 ---LRGRLDLSNKNLVM---QTLTTIGGVAAMPAPAEKASKGILLSVLKCLGDNKCKMR 959
Qy 689 IKSLA-----FGPDRIEMLEPEVL-----DGRM-----LKIMTDRAETAQGLVYDYT 732
Db 960 ECTLAALDLMTGAVHLDMMITPIYIIALTLDGKMGABGRKDLDMW-----TKQLTGLSDPV 1014
Qy 733 ---LKLTVSDSVTVSADGTRALVEATLEBSACLSIDLVPENNAVDV 776
Db 1015 DAHTLLKPASTAMTKSADVAKAEGCISILRVSQGMIEKNLKDIO 1062
RESULT 7
Db 732
DNA-directed DNA polymerase (EC 2.7.7.7) - human herpesvirus 3
C:Species: human herpesvirus 3, varicella-zoster virus
C>Date: 30-Sep-1998 #sequence_revision 30-Sep-1998 #text_change 09-Jul-2004
C/Accession: B27214
R/Davison, A.J.; Scott, J.B.
J. Gen. Virol. 67, 1759-1816, 1986
A>Title: The complete DNA sequence of varicella-zoster virus.
A/Reference number: A27345; MUID:86306657; PMID:3018124
A/Accession: B27214
A/Molecule type: DNA
A/Residues: 1-1194 <DAV>
A/Cross-references: UNIPROT:P09252; EMBL:X04370; NID:g59989; PIDN:CAA27911.1; PID:g60017
C/Genetics:
A/Suprafamily: DNA polymerase
C/Keywords: DNA binding; DNA replication; nucleotidyltransferase
Query Match 3.2%; Score 131.5; DB 1; Length 1194;
Best Local Similarity 19.0%; Pred. No. 4.1;
Matches 166; Conservative 122; Mismatches 297; Indels 297; Gaps 44;
Qy 48 SDNFNTSDSS---SSPATATTATLVSLPSIDRPERHV---PIPIDYQVLAQ---TH 99
Db 205 NDAITLNDKNAFHGTSFKASPSFSEV---ERTDYVYDTQPCAFYRVSPSSKFTN 261
Qy 100 FLVDGIR---RAFAVRSKPPQGFSDALIS-----RQI----- 132
Db 262 YLCDNFHPELKKCEGRDATTFRFLMDNPGVFSFGWYQLKPGVDERVAVRPASHQTLSD 321
Qy 133 LQAAAC-----ETLSNPSRRBYNEGLDDR-----BATVITVDVWMDKVPALC 175
Db 322 VEIDCSDNLOALPNDSSWPDYKLLCFDIECKSGSNBLAPDITHELDL---VIOISC 377
Qy 176 VL---QEGGETEIVLRVGEALLKRLPKSEFKQ-----DVLVMAALFDVSRD 220
Db 378 LTVSIRPQSLHEHLLFELGSCDLPQRYVQEWKQAGLPFVLPBDSPELLIAFMVLVKQ 437
Qy 221 AMALDPDPFITGTFEVEALKLQEBGASSLAPDLRAQIDETLEITPRYVLELGLPLG 280
Db 438 Y---AEEFATGVYIVFNAFI-----MEKLSISYLSKI----- 468
Qy 281 DDVAAKRLNGISGRNLTMSVGGGASALVCGLTRBKPMNEAFRMTAAEQVDLFAVTPS 340
Db 469 DGGGSIIRGGLFKI---WDVKGSG-----FQRRSKVYKINGLSLDVMAIATE 512
Qy 341 NIPASEFEVEVA-----LATVAQAFITG 363
Db 513 KTKLSYKLDVSVAEALNESKRDLPYKIDPGYVAGSPNTRGIIIGEYCIOBSALVGLKFPK 572
Qy 364 KKHLLDADKQFQOLQAKVMAMEIRPMLYDTRNNEIDIGLERGLCALLIGVDECRM 423
Db 573 YLPLELSA-----VARLARITLTKAYD-----GOQVRYTYCIIAGASS--- 612
Qy 424 WLGLSDSDSOYRNPAIVEF-----VLENSNRDDNDLPGCKLLETVLAVGV-VEPR 473
Db 613 -RGITLDPGGY--PATEYKQVIDVGDVBEKMBDESVSPYSGSNVYKQARVPD- 668
Qy 474 FRDTKKKFKLGDYDDPMVLSTYLERV-EVVGSGPLAAATMAIRGAHVVASMAQLQK 532
Db 669 -PDI-----GFYIDPVVVLDPFASIVPSIIQAHNL--CFTTLTINFETVK-----R 710

Qy 533 VFPERYTDRNSABPKOVETVFSYDP-----VGNNGRDEPGVFAE--AVRESENET 585
Db 711 LNPEDY-----ATFTYVGKRLFFVRNSV-RESLITGLLMDMLMRK----- 750
Qy 586 NDVAIRAGVSESSVDEFTVENSVAIDMEASVKILA-----AGVAIGLIS----- 630
Db 751 ---AIRARIPSSSDE-----AVLLDKQAIAIKVNCVSGFTVAGCFPLCYVAATVT 802
Qy 631 -----LFGQKYFLKSSSGFR-----KQVVSMSGDVA-TISGVADDEALPRMDAR 677
Db 803 TIGRQMLSTRDYIHNNMAAPERFITAPPDIESSVLSQKAVEVYIGDTSVFIREKGV 862
Qy 678 TAENIVSKMOK-----IKSLAFGPDRIEMLEPEVLGKIMTWRRAETAQGLVYDYTL 733
Db 863 SVEGIATIGEMMAHITSTALFCPPIKICEKTFI--KLILI-----TKKKYIGVYGGKV 915
Qy 734 LKLSVDSVTVS---ADGTRALVEATL-----ESBACLS 763
Db 916 LMKGVDLVRKNNCOFINDYARKLYELLDYDPTVSRRAAASCVS 959
RESULT 8
A70668
mycobacteroate synthase (EC 2.3.1.111) - Mycobacterium tuberculosis
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: A70668
R/Cole, S.T.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S. Nature 393, 537-544, 1998
A/Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: A70668
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-2111 <COL>
A/Cross-references: UNIPROT:P96291; GB:Z83858; GB:AL123456; NID:g3261675; PIDN:CAB06108.1
A/Experimental source: strain H37Rv
C/Genetics:
A/Suprafamily: mycobacteroate acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I hom
nase homology; [acyl-carrier-protein] S-malonyltransferase homology
C/Keywords: acyltransferase; carrier protein; coenzyme A; phosphotransferase
F/27-426/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F/534-815/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F/1444-1733/Domain: long-chain alcohol dehydrogenase homology <LADH>
F/1765-1945/Domain: short-chain alcohol dehydrogenase homology <SADH>
F/2016-2096/Domain: acyl carrier protein homology <ACPI>
Query Match 3.2%; Score 131.5; DB 2; Length 2111;
Best Local Similarity 19.9%; Pred. No. 10;
Matches 151; Conservative 82; Mismatches 237; Indels 289; Gaps 33;
Qy 151 EGLIDSEATVITDP---NDK-----VPALCYLQEGGETEIVLRVGEALLKRL 198
Db 29 ESLIRGDD--LVTEIPDRDADDDYDPEPGVGR-SVSRGQ----- 68
Qy 199 PKSPKQDVVLMALAFDVSRDMAALDPDPFI---TGVEFEELAKLLQEBGASSLAPDL 255
Db 69 ---FLDDVAGFPAEFQISREARSTIPQGRLLLETWEALIEH----- 109
Qy 256 RAQIDETLEITPRYVLELGLPLGDDVAAKRLNGLSGVRNLTMSVGGGASALVGLTR 315
Db 110 -----GLDPASL-----AGSTAVFTGLTH 129
Qy 316 EKPMNEAFLMNTAAEQVDLVATPSNI-----PASEFEV----- 350
Db 130 EDYL---VLTITGGLASPYVVTGLNNSVASGRIATHTLGLHGPMTDPTDTCSSGLMAVHL 186
Qy 351 -----EVALALVAQAFIGKKPHLL-----ODADKQFQOLQAK 383

```
Db 187 ACSRLHDEADLLAGCAVLEBPHASVAASAOGLMSTGRCHLFDADAGFVNSGCM 246
QY 384 VMAMEIIPAMLYDRNNWEIDFGLERGLCALLIGKVECRMWLGSDSGYRNPATVEVF 443
Db 247 VLKRLPDLARDGR-----IFAVRGTAITNODGRTER-----LTPSBDAYV--AVRAA 295
QY 444 LENSNRDDNDLPGCLKLETLWAGVFPFRFDTKDKKFKLGDDYDDPMWLSYLERVEV 503
Db 296 LAAG-----VQBEIVGV-----BAHGTEGPIGDPIEYRSIARV-YG 332
QY 504 QGSPPLAAATMARIIGAHHVKSAMQALQKV-----PRTYDRMSAPKQVQVEVFS 555
Db 333 AGTPCALGSAKSNWGSTASAGTGLIKALISLRHGVPPPLHFNRLPDELSDV----- 386
QY 556 VDPVGNVVGDEBPGVIAEAVRSENFETNDVA-IRAGVSESVDTEVWMSYADMLKE 614
Db 387 -----ETGLFVPAQVTPWPN--GNDHTPKVANVSSFGMSGTNHAYVEAPAE 432
QY 615 ASVKIILAAVVA-IGLISLFSQKVFLLKSSSFQKDMVSSMESDVATIGSVRADSEALPR 673
Db 433 ASAPESPQDAEYV-----PRLFPLMSTSS-----SDAL-R 461
QY 674 MDRPTENITYSKQK-----IKSLAFGPDHR-----IEMLPETLDGRML 712
Db 462 QTRAROLATWVEEHQDCVAASDLAYTLARGRAHREVRTAVVAANLPELVEGLREVADGAL 521
QY 713 -----KIW--TDRAETPAQLGL-----VYDTYLLKLSVDSVTVSVADGTRALV 752
Db 522 YDAVAGHGBRGFPWTVVSGGSSQWAAKMTQLLASPEVPAATIAKL---EPVIAAESGFSVT 578
QY 753 EA-TLESACLSDLVHPENNATDVRYTYTREVFWMSKSG 790
Db 579 EAITAQOTVVGIDKVPAPVAVQVALAATMEQFGVARG 617
```

RESULT 9

```
C97783
cell surface antigen [Imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C:Date: 30-Sep-2001 #sequence_rev1sion 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: C97783
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rd
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: C97783
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1026 <KUR>
A:Cross-references: GB:AE006914; PIDN:AL03205.1; PID:G15619756; GSPDB:GN00173
C:Genetics:
A:Gene: sca4
```

```
Query Match 3.24; Score 130.5; DB 2; Length 1026;
Best Local Similarity 18.84; Pred. No. 3.71;
Matches 124; Conservative 109; Mismatches 277; Indels 151; Gaps 25;

QY 115 KPPQFESDIALSRQIILQACETLSNPSRBRVNGGLDDEBATYITVPMKV----- 170
Db 378 EPPQYKQVPPPIPTPQPIQP--ETSQMPOS-QQVNNPLNLN--TATLSSGMDLMLVYN 432
QY 171 PGALCVLQEGGETEIVLRVGEALLKERLPSKFKQDVVLVNA--LAFDVRSDAMALDPPD 228
Db 433 AGITKALIDSKKQDLIKREATAILNNE-----KSDIAEKQANIILAEVTVNNKMLKPA 487
QY 229 FITGYEFVEBALKULOEGGASSLAPDIRAOI-----DETLEETPRVYEL---LGLPLG 280
Db 488 KVTGVAVAVETIKRDQNTPMLEKSKMLEATVAIVLSENLEPKQKQOMLEKAVDVGSLK 547
QY 281 DQVAAKRLNGSLGVRNITLWVGSGGASALVGLTRFENMEAPLRMAAQQVULFVATPS 340
Db 548 DD--ASRAAAIDIGIKDVV-----IKSNLSPEDKKMLAVGD 580
```

```
QY 341 NIPAESFEVYEVALVAQAFIGKKPHLLQDADKQFOQLQOAKVMA-----MEIPAM 392
Db 581 KVVVSELSNAEKQKLIGSVLKKEVEAQVLSPPAQOQLMQHLYYIMAEQTKDITIKVNDI 640
QY 393 LYDTRNNWE-----IDFGLERGLCALLIGKVECRMWLGSDSGYRNPATVEVF 439
Db 641 LFDPPLSTBELKTTNIOAITSVNLDGPATVAEKKELLQAITNTTAGSSLEAD---KAAL 696
QY 440 VEVYLEN--SNRDNDLPGCLKLETLWAGVFPFRFDTKDKKFKLGDDYDDPMWLSYL 497
Db 697 IKVGGERIAHSDTSLSLPKALIMASAEKGI-----EQTNLPD--RELMTGLV 746
QY 498 ERVEVQGSPLAAATMARIIGAHHVKSAMQALQKV-----PSSRYTDRN 542
Db 747 DGIVEGKGBEITKAVSSGIDNSININDSEKALKKKAADAABALDRDTQNLTEGFGQN 806
QY 543 SAKPK-----DVQEVFSVDVPVGNVNGDGPVFIIE-----AVR 579
Db 807 IEERKPHDDIYNKAREVINAVNPIEALSKSPVSAEERIVQETSSILNNISKLAVEK 866
QY 580 SENFETNDVAIRAGVSESVDTEVWMSYADMLKEASVKIILAAVVAIGLISLFSQKVFLLK 639
Db 867 VNNF-----PAMLSPNGMLKTLEEK-----KEBAIKRVDE-----LYKAGTGSSTE 908
QY 640 SSSSFQKDMVSS--SMESDVATIGSVRAD--SEALPMDARLYAE-NIVSRKQIKISL 692
Db 909 EQGSFIKTNLIDDKTSLKSVRLQITDKLQEQKRSKAIENPVSATBEDVRVVGSKSLKPI 968
QY 693 A 693
Db 969 S 969
```

RESULT 10

```
B44110
mycoceroate synthase (EC 2.3.1.111) - Mycobacterium bovis
C:Species: Mycobacterium bovis
C:Date: 05-Dec-1998 #sequence_rev1sion 05-Dec-1998 #text_change 09-Jul-2004
C:Accession: B44110
R:Mathur, M.; Koliatukudy, P.E.
J. Biol. Chem. 267, 19386-19395, 1992
A:Title: Molecular cloning and sequencing of the gene for mycoceroate acid synthase, a nc
-Guerrin.
A:Reference number: A44110; MUID:92406887; PMID:1527058
A:Accession: B44110
A:Molecule type: DNA; protein
A:Residues: 1-2110 <MAT>
A:Cross-references: UNIPROT:Q02251; GB:M95808; NID:G149977; PIDN:AAA25369.1; PID:G149980
A:Note: Sequence extracted from NCBI backbone (NCBIN:113897, NCBI:P:113899)
A:Note: the source is designated as Mycobacterium tuberculosis var. bovis Bacillus Calmet
C:Superfamily: mycoceroate acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I hom
nase homology; [acyl-carrier-protein] S-malonyltransferase homology
C:Keywords: acyltransferase; carrier protein; coenzyme A; phosphotransferase
F:27-426/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
F:534-815/Domain: [acyl-carrier-protein] synthase I homology <OAS>
F:1443-1732/Domain: long-chain alcohol dehydrogenase homology <LADH>
F:1764-1944/Domain: short-chain alcohol dehydrogenase homology <SADH>
F:2015-2095/Domain: acyl carrier protein homology <ACP1>
```

```
Query Match 3.24; Score 129.5; DB 2; Length 2110;
Best Local Similarity 19.94; Pred. No. 13;
Matches 151; Conservative 82; Mismatches 237; Indels 289; Gaps 33;

QY 151 EGLLDDEBATVITDVP---WDK-----VPGALCVLQEGGETEIVLRVGEALLKERL 198
Db 29 ESLRGDD--LVTEIPDRKRDADYDDEPGVGR-SVSMWG----- 68
QY 199 PKSFQDQVVLVMAIAFLDVRSDAMALDPPDFI---TGYEFVEBALKULOEGGASSLAPDL 255
Db 69 ---FLDDVAGFADFEGISREARATSIDPQORLLLETSWEAIEHA----- 109
QY 256 RAOIDETLEETPRVYELIGLPLGDDYVAAKRLNGSLGVRNITLWVGSGGASALVGLTR 315
```

Db 110 -----GLDPASL-----AGSTAVFTGLTH 129
 Qy 316 EKPMNEAFLRNTAAEQVDLFPVATPSNT-----PASPEVY----- 350
 Db 130 EDYL---VLTTTAGGLASPYVVTGLNNSVASGR1AHTLGLHGPATPTACSSGLMAVHL 186
 Qy 351 -----EVALVALVAQAFIGKPPHL-----ODADKQFOOLQOAK 383
 Db 187 AGRSLHGERDLALAGCAVLLEBPACVAAQAQMSLSTGRCHSDADADAFVBSGGAM 246
 Qy 384 VMAMEIPAMLVDTNNNEIDFGLERGLCALLIGKVDCEKMWIGLDSEDSQYRNPAIVEFV 443
 Db 247 VILKRLDRLADRGSR-----IFAVRGRTATNODGRRET-----LTPMSBAQV---AVYRAA 295
 Qy 444 LENSNRDDNDLPLGLCKLETLWLAGVFPFRPRDTKDKFKLGDYDDPMVLSYERREVV 503
 Db 296 LAAG-----VQPEVTVGVV-----EAMGTGTPIGPIEYRSIARV-YG 332
 Qy 504 QGSPLAAATMARIAGAHVKSAMQALOKVF-----PSRYTDRNSAEPKDVOETVPS 555
 Db 333 AGTCALGSAKSNHGSTASAGTGLIKAILSLRHGVVPLHLFRRLPDELSDV----- 386
 Qy 556 VDPVGNVNGDEBPGVFIAEAVRPSSENEETNDYA-IRAGVSESSVDTEVMSVADMLKE 614
 Db 387 -----ETGLFVPQAVTWPMPN--GNDHTPKRAVAVSSFGMSGTNVAHIVEAPAE 432
 Qy 615 ASVKIILAGVA-IGLISLFSOKYFLKSSSSFORKDMVSSMESDVATTGSGRADSEALPR 673
 Db 433 ASAPSSPGDAEVG-----PRLFMLSSTS-----SDAL-R 461
 Qy 674 MDARTANIVSKWK-----IKSLAFGPDHR-----IEMLEPYLDGBML 712
 Db 462 QTRAQALATWEEHODCCAASDLATLARGRRHRYRAVAANPELVESGLREVAADGAL 521
 Qy 713 -----KIW--TDRAETALGL-----VYDTLLKLSVDSVTVSADGTRALV 752
 Db 522 YDAVGHGDRGPVWVFGSGSQSOMAMGTOILLASEPFAATIAKL---EPVIAASGSPSVT 578
 Qy 753 EA-TLESACISLIVHPENNATDVRTYTTREYVFWMSKSG 790
 Db 579 EATTAQCTVIGIDKVPVAVFVQVALATWETQYGVHFG 617

RESULT 11
 T49362
 hypothetical protein Bld1.160 [imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
 C:Accession: T49362
 R:Schulte, U.; Align, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000
 A:Reference number: 225022
 A:Accession: T49362
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1275 <SCH>
 A:Cross-references: UNIPROT:Q996A4; EMBL:AL355927; GSPDB:GN00116; NCSP:Bld1.160
 A:Experimental source: BAC clone Bld1; strain OR74A
 C:Genetics:
 A:Gene: NCSP:Bld1.160
 A:Map position: 6
 A:Introns: 24/3

Query Match 3.2%; Score 129; DB 2; Length 1275;
 Best Local Similarity 20.9%; Pred. No. 6.5;
 Matches 114; Conservative 84; Mismatches 210; Indels 138; Gaps 26;

Qy 325 RMTAAEQVDLFPVATPSNIPASFEVYEVALLVAQAFIGKPPHL----- 368
 Db 39 RMTFATPTAVFAALLSLNNENLATPVVYLLDVVTPF-APRPILIQAKFTQITITLAPLVS 97
 Qy 369 LGADKQFOOLQOAKVMAEIPAMLVDTNNNEI---DFGLERGLCALLIGKVDCEKMWI 425

Db 98 LPDAD---APLMRASISGLE-SILLADATQWEIGTAQIGRRVAVAGLLA-----L 144
 Qy 426 GLDESDS-QYRNPAIVEVLENSNRDNDLPLGLCKLETLWLAGVFPFRDTKD----- 479
 Db 145 ALDRPVPKRAQBALNRILNPPSPSLDHPAPMCAETAMQSL-----RLVADKAAQ 199
 Qy 480 -KFKLGDYDDPMVLSYERREVVQ-----GSPLAATM-----ARIGAEHVKSAM 527
 Db 200 RKEKTTDSTDPDLIALQLVKAVASGGGWPKSLSLCELLSTARTGNHMSAVF 259
 Qy 528 QALQKVEPSRYTDR-NSAEPKDVOETVFSVDPVGN-----VGRDEPGVFIAEA 576
 Db 260 EIFEMFEGMAADEVASAKPLRLLEIRLELPAPNDQLLEPWIALLSRAVD-----VAGQ 315
 Qy 577 VRSENE--TNDYAIAGVSES-----SVDETTEMSVADMLKERSV-----KIL 620
 Db 316 ISPAETFOELVDPFTLVAGYLSQHKNIIRVSASCLVSPLANCVPKKALIEPSIFDERVI 375
 Qy 621 AAGVAI---GLISLFSOKYFLKS-----SSSFORKDMVSSMESDVATTGSGRADSE 669
 Db 376 QQLVKVVEGLITVOYQAAMETFNVLGAMPDAPRWQANPYLLSV---VKSIGELRGDSF 432
 Qy 670 ALPMDARTANIVSKWKIKSLAFGPDHRIEMLP-----EVLDRMLKIWTDRAAETAO 725
 Db 433 A---GKQEADEVLGKAIR---AMGPDAVLNVPLMLAKVKGQPGRAW-----ML 476
 Qy 726 GLVDTY---LKLSDSVTSVSGDTRALVEATLESACISLIVHPENNATDVRTYTT 781
 Db 477 PLLEDTYSNTNLHAFKSELVPLSAIMFORVIE-----HOGKNIHIKIFETV 523
 Qy 782 YEVEFWS 787
 Db 524 VQQTWS 529

RESULT 12
 T30312
 pilin biosynthetic protein - Pseudomonas aeruginosa
 C:Species: Pseudomonas aeruginosa
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 04-Mar-2000
 C:Accession: T30312
 R:Whitchurch, C.B.; Young, M.D.; Hobbs, M.; Mattick, J.S. submitted to the EMBL Data Library, November 1996
 A:Description: Pseudomonas aeruginosa chemotactic transduction genes pilL, cnpA cnpB and
 A:Reference number: 220819
 A:Accession: T30312
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-836 <WHI>
 A:Cross-references: EMBL:U79580; NID:g3241967; PID:g3241968; PIDN:AMC23930.1
 C:Genetics:
 A:Gene: pilL
 C:Superfamily: Pseudomonas aeruginosa pilin biosynthetic protein

Query Match 3.2%; Score 128.5; DB 2; Length 836;
 Best Local Similarity 20.0%; Pred. No. 3.6;
 Matches 158; Conservative 106; Mismatches 292; Indels 235; Gaps 35;

Qy 124 DALISRRQIIQAACETLSNPRSRREYNEGLDDEAATVITDVPDKYFGALCVLQEGGE- 182
 Db 24 ETLKQARQALEAFENQDPTRKR-----FCLTVV--HVGGLTQWMEFVGAA 69
 Qy 183 --TSIVLRVGBALIKERLPKSPKODVVLVMAFLVSRDAMALPPDF-----ITGYE 234
 Db 70 LLAEMEQALVOLALDGGVPPVNGEALBVLMOAIIQLPYLDRICQFARBDLPVVVLPILNDL 129
 Qy 235 FVEBALKLIOEBGSSLAAPL---RAQID-----ETLEETTPRYVLELGLP 278
 Db 130 RAARGELKLSB--TSLAPDLSQRQPDLDGEALAQLTDEGLGLKRLKROTQOMALVGLL 187
 Qy 279 LGDDVAAK-----RLNGLSGVRNI--LWSVGGGASALVGLLTREKPMNEAFLRMT 327

```

Db      188 RNDVATSLGLAVYARLEGLCREAPGLMSI-----ASGLVEGLANGSVNSASVR-T 242
Qy      328 AAEVD-----LFAVTPSN-----PAEPEVEVALVAQAFIG 363
Db      243 LIRQJDEBELKRYEQAGDLNQAAPDELYKNLRYAKAPSQSPRIRAL----- 291
Qy      364 KKPRLQDA--DKQFOOLQAKVAMEIPAM-----LYDTRNWEIDFGLERG-- 409
Db      292 KEQYRLDEALPDHETVAERARLAGPRDANRSVVALCEELVRIKQSLDFVNSDGHPR 351
Qy      410 --LCALL--IGKUDCEKMLGLDSEDSQYRNPAIYERFLSNSRDDND---DLPGCLK 461
Db      352 SEDLLALAPKQJLDLAVLGFQGRKRVYILDQDLVHIALAQGRREPSDAILMDVAGALLY 411
Qy      462 LETLAGVFP--RPDTRDKKKFLGDYD-----DPMVLSYL----- 497
Db      412 VEATLAGMAGBGRNBERSESLPTTVDVAQHQLVYKERRNGLEQAKAIIEFIFASQNMHE 471
Qy      498 -----ERVEVQGS-----PLAAATMARIGAEHVKASAMQALQKVP----- 535
Db      472 HLAARPELLQVRRGLAMIFLERAAATLLETGNRYIQEQL--ARKAVPDMQSLDTLDAI 529
Qy      536 ---SRYDRNSAEKXVOEYFVS--DPVGN-----NCRDGEFGVFLAEAVRSENF 583
Db      530 TSEYTYLERLSEDAQSQDLILDVAEDSLANIGYTLKXNSSAPAPFGI----- 577
Qy      584 ETNDYAIRAGVSESSVDETTVMGVADMLKEASVKILAAAGVAILISLFSQKFLKSSS 643
Db      578 -SGPAALESRAEERPEAVE--VAETAQPPADTPALEA----- 616
Qy      644 FORKDWVSMESDVATIGSVRADSEALPRMDARTAEINYSKWKIKSLAFGPDRIEML 703
Db      617 --BEDAPQLASDDWMTIGEVAPDAGE--PSLDL-----ALDLPDDSAEVP 658
Qy      704 PEVLDMGLKIWDRAETQGLGVYTYTLKLSVDSVTYSAD-----GRALVEATLEE 758
Db      659 PAVEECRE---TPPOSTSAPARSLDFSLDEIDLSGLDLPADAAPASGPAALADWLSLPR 714
Qy      759 SACL--SDLVHP 768
Db      715 QMGLGDDDLAQ 725

RESULT 13
S70368
DNA polymerase I - Bacillus stearothermophilus
C:/Species: Bacillus stearothermophilus
C:/Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:/Accession: S70368
R/Rig98, M.G.; Tudor, S.; Sivaram, M.; McDonough, S.H.
Biochim. Biophys. Acta 1307, 178-186, 1996
A:/Title: Construction of single amino acid substitution mutants of cloned Bacillus stear
A:/Reference number: S70368; MUID:96283813; PMID:8679703
A:/Accession: S70368
A:/Status: preliminary
A:/Molecule type: DNA
A:/Residues: 1-876 <RIG>
A:/Cross-references: UNIPROT:P52026; EMBL:L42111; NID:9806280; PIDN:AA37139.1; PID:98062
C:/Genetics:
A:/Start codon: TTG
C:/Superfamily: DNA-directed DNA polymerase I

Query Match      3.2%; Score 128.5; DB 2; Length 876;
Best Local Similarity 21.0%; Pred. No. 3.9;
Matches 170; Conservative 135; Mismatches 302; Indels 201; Gaps 43;

Qy      99 HFTLDGIRAFEARVSKPQFGSDALISRRQILQACETLSNPSRRRY-NEGLLDE 157
Db      107 HYADADILIGMAARARE--GFAVKVISGDRDLQ-----LASPVVETITKKGIRDIE 158
Qy      158 EATVIVDY-----PMDKVRG-----ALCVQEGGTEIYILRVG 190
Db      159 SYTPETVVEKYGLTPQIVDLKGLMGDKSDNIPGVGIGEKTYAVKLKQFGVENVYLA 218

```

```

Qy      191 EALLKERLPKSFKO--DVLVLMALAFADVGRDA---MALDPDFTGVEFEELKLLQE 245
Db      219 DEIKGEKLENLQYRDIAL-LSKQLAICRDAPEVELTD--DIVYKGEERKYNVLFQE 275
Qy      246 EGASSLAPDLRAQID-----TLEETPRVYLLLEIG--PLGDDYAAKRLNG- 290
Db      276 LGFQSPDKMAVQTDDEGEKPLAGMDFAIDSVYDEMADKAAAVVEVVGNDYHNAPIVGI 335
Qy      291 -----LSGVRLNLSVGGGGAALVQGLTREFENMEAFIEMTAEE----- 330
Db      336 ALANERGFRLREYALADKFLAW-----LDETEYKKTMDSKAAVALKMKGI 385
Qy      331 ----QVDFVATPSPNIPAES-----PEVEVALALVAQAFIG-----KPHLQ 370
Db      386 ELRGVFDLLAAVLDPRQAAGDVAAVAMGHYEAVRSEAVYGGAKARTVDEPTLAE 445
Qy      371 DADKFOQLQQAQKVMAMEIPAMLYDTRNWE-IDFGLERGLCALLG-----KVDECRM 423
Db      446 HLAAR-----KAAAIWALEEPLMDLRNNEQDRLLTELEQPLAGIILANMEFTGVKVDTKRL 500
Qy      424 -WIGLD-SEDSQYRNPAIYERFLSNSRDDNDLPGCLKLEFWLAGVPRRRTDKDK 481
Db      501 EQWGAELTEQLQ---AVERRIYELAQEENINSP---KQLGT---VLPDKLQLPVLKX 549
Qy      482 FKLGDYDDPMVLSYL--ERVEVQGSPLAAATMARIGAEHVKASAMQALQKVPSPRYTD 540
Db      550 TKTG-YTSTADVLEKLAHPHEIYE--HILHYRDLGLQSTYIE-GLLKVNHPYTGKVNHTM 605
Qy      541 RNSAEPKVOET--VPSVDVGNNGRDEGPGVFLAEAVRSE-----NFTNDYAI 590
Db      606 FNOA---LTQTRLSVSEPNLQNIPIRLIEGGRKIQAEFVPSBPDWLIFADYSQJELRV 661
Qy      591 RAGVSESSVDETFE-----MSVADMLKE---ASVKILAAAGVAILISLFSQ 634
Db      662 LAHIAE--DDNLEAFRRGLDHTKTAADIHFVSESDVTANRQAKANFGIIVGISD 718
Qy      635 KYFLKSSSFORKDWVSMESDVATIGSVRADSEALPRMDARTAEINYSKWKIKSLAF 694
Db      719 -YGLAQNLNITRKEAAEFIRYFASFPQVK-----QYMDNIYGE-AKQGYVT 764
Qy      695 GPDRIEMLPEVLDMGLKIWDRAA-ETQAGLVYDYTLKLSVD-SVTSADGTRA- 750
Db      765 TLHRRRYLPDIRSRFNVSFAERTAMTPPIQSAAD-IKKAMIDLVSRLREERLQAR 823
Qy      751 -----LVEATLESACLSDLVHPE 769
Db      824 LLLQVHDELLEAPKEIERLCRLV-PE 850

RESULT 14
E83594
Still frameshift probable component of chemotactic signal transduction system PA0413 [lm]
C:/Species: Pseudomonas aeruginosa
C:/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:/Accession: E83594
R/Stover, C.K.; Pham, X.Q.; Ewitt, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri
aden, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Larbig, K.; Llm,
.; Lory, S.; Olsen, M.V.
Nature 406, 959-964, 2000
A:/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathog
A:/Reference number: A82950; MUID:20437337; PMID:10984043
A:/Accession: E83594
A:/Status: preliminary
A:/Molecule type: DNA
A:/Residues: 1-2472 <STO>
A:/Cross-references: UNIPROT:Q91696; GB:AE004478; GB:AE004091; NID:99946261; PIDN:AA60380;
A:/Experimental source: strain PA01
C:/Genetics:
A:/Gene: PA0413

Query Match      3.1%; Score 127.5; DB 2; Length 2472;
Best Local Similarity 20.0%; Pred. No. 23;

```


THIS PAGE BLANK (USPTO)

QY 361 FIGKKPHLLQDADKQFOLOQAQKVMAMEIPAMLYDTRNNMEIDFGLERGLCALLIGKVD 420
 DB 361 FIGKKPHLLQDADKQFOLOQAQKVMAMEIPAMLYDTRNNMEIDFGLERGLCALLIGKVD 420
 QY 421 CRWMLGDSDSQYRNPAIVEFVLENSNRDNDLPGLCKLETWLAGVPPRRDTKDX 480
 DB 421 CRWMLGDSDSQYRNPAIVEFVLENSNRDNDLPGLCKLETWLAGVPPRRDTKDX 480
 QY 481 KFKLGDYDDPMTLSYLERVEVQSSPLAAATMARI GAHVKASAMQALQKVPSSRYTD 540
 DB 481 KFKLGDYDDPMTLSYLERVEVQSSPLAAATMARI GAHVKASAMQALQKVPSSRYTD 540
 QY 541 RNSAEPDQVEITFSVDPVGNNGRDEPGVFIAEAVRPSSENFETNDYAIRAGVSSSD 600
 DB 541 RNSAEPDQVEITFSVDPVGNNGRDEPGVFIAEAVRPSSENFETNDYAIRAGVSSSD 600
 QY 601 ETTVEKSVADMKEASVKIIAAGVAIGLISFQKYLKSSSFQRKDMVSMESDVATI 660
 DB 601 ETTVEKSVADMKEASVKIIAAGVAIGLISFQKYLKSSSFQRKDMVSMESDVATI 660
 QY 661 GSVRADSEALPRMDARTAEINYSKMOKISLAFGPHRIEMLEPVLDGRLKIWDRAA 720
 DB 661 GSVRADSEALPRMDARTAEINYSKMOKISLAFGPHRIEMLEPVLDGRLKIWDRAA 720
 QY 721 ETKQLGLVYDYLKLSVDSVTSVADGTRALVEATLEESACLSLVHPENNADVTYTT 780
 DB 721 ETKQLGLVYDYLKLSVDSVTSVADGTRALVEATLEESACLSLVHPENNADVTYTT 780
 QY 781 RYEVFWSKSGMKITEGSLAS 801
 DB 781 RYEVFWSKSGMKITEGSLAS 801

RESULT 2

QYXAR9 ID QYXAR9 PRELIMINARY; PRT; 801 AA.
 AC QYXAR9; DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Division protein.
 GN Name=AKC6;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22779057; PubMed=12897262;
 RA Vtcha S., Froehlich J.E., Kosharova O., Pyke K.A., Van Epp H.,
 RA Osteryoung K.W.;
 RT "ARCE is a J-domain plastid division protein and an evolutionary
 RT descendant of the cyanobacterial cell division protein Ftn2.";
 RL Plant Cell 15:1918-1933(2003).
 DR EMBL; AY221469; AAC18646.1; -.
 DR InterPro; IPR001623; DnaJ N.
 SQ SEQUENCE 801 AA; 88247 MW; 7E2E1B3FD4B61 CRC64;
 Query Match 99.7%; Score 4052; DB 2; Length 801;
 Best Local Similarity 99.8%; Pred. No. 3.3e-240;
 Matches 799; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 121 FSDALISRRQIIQAACETLSNPRSREYNELGLIDDEGATVITDVPMDKVPALCVLQEG 180
 QY 181 GETEIVLAVGEALLKELPKSPKQDVVLYMALFLDVSRDAMALDPPDITGYEVEEAL 240
 DB 181 GETEIVLAVGEALLKELPKSPKQDVVLYMALFLDVSRDAMALDPPDITGYEVEEAL 240
 QY 241 KLIQEBGASSLAPDLRAQIDETLEETPRVYELGLPLGDDYAAKRLNGSGVRITMS 300
 DB 241 KLIQEBGASSLAPDLRAQIDETLEETPRVYELGLPLGDDYAAKRLNGSGVRITMS 300
 QY 301 KLIQEBGASSLAPDLRAQIDETLEETPRVYELGLPLGDDYAAKRLNGSGVRITMS 360
 DB 301 KLIQEBGASSLAPDLRAQIDETLEETPRVYELGLPLGDDYAAKRLNGSGVRITMS 360
 QY 361 FIGKKPHLLQDADKQFOLOQAQKVMAMEIPAMLYDTRNNMEIDFGLERGLCALLIGKVD 420
 DB 361 FIGKKPHLLQDADKQFOLOQAQKVMAMEIPAMLYDTRNNMEIDFGLERGLCALLIGKVD 420
 QY 421 CRWMLGDSDSQYRNPAIVEFVLENSNRDNDLPGLCKLETWLAGVPPRRDTKDX 480
 DB 421 CRWMLGDSDSQYRNPAIVEFVLENSNRDNDLPGLCKLETWLAGVPPRRDTKDX 480
 QY 481 KFKLGDYDDPMTLSYLERVEVQSSPLAAATMARI GAHVKASAMQALQKVPSSRYTD 540
 DB 481 KFKLGDYDDPMTLSYLERVEVQSSPLAAATMARI GAHVKASAMQALQKVPSSRYTD 540
 QY 541 RNSAEPDQVEITFSVDPVGNNGRDEPGVFIAEAVRPSSENFETNDYAIRAGVSSSD 600
 DB 541 RNSAEPDQVEITFSVDPVGNNGRDEPGVFIAEAVRPSSENFETNDYAIRAGVSSSD 600
 QY 601 ETTVEKSVADMKEASVKIIAAGVAIGLISFQKYLKSSSFQRKDMVSMESDVATI 660
 DB 601 ETTVEKSVADMKEASVKIIAAGVAIGLISFQKYLKSSSFQRKDMVSMESDVATI 660
 QY 661 GSVRADSEALPRMDARTAEINYSKMOKISLAFGPHRIEMLEPVLDGRLKIWDRAA 720
 DB 661 GSVRADSEALPRMDARTAEINYSKMOKISLAFGPHRIEMLEPVLDGRLKIWDRAA 720
 QY 721 ETKQLGLVYDYLKLSVDSVTSVADGTRALVEATLEESACLSLVHPENNADVTYTT 780
 DB 721 ETKQLGLVYDYLKLSVDSVTSVADGTRALVEATLEESACLSLVHPENNADVTYTT 780
 QY 781 RYEVFWSKSGMKITEGSLAS 801
 DB 781 RYEVFWSKSGMKITEGSLAS 801

RESULT 3

QYFIC9 ID QYFIC9 PRELIMINARY; PRT; 801 AA.
 AC QYFIC9; DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone:MDH9
 DE (Hypothetical protein At5g42480).
 GN Name=At5g42480;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93156233; PubMed=10048486;
 RA Aamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned pl and TAC clones.";
 RL DNA Res. 5:379-391(1998).
 RN [2]
 RP SEQUENCE FROM N.A.

```
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Dang J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shim P., Southwick A., Shinzaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan Y.W., Lee J.M.,
RA Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shim P.,
RA Southwick A., Trapp M.G., Wu T., Shinzaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB016888; BAB10489.1; -.
DR EMBL: AY091075; AAM13895.1; -.
DR EMBL: AY150490; AAN12907.1; -.
DR InterPro: IPR001623; DnaJ_N.
DR Hypothetical protein.
SQ SEQUENCE 801 AA; 88259 MW; 608E776FBA73FECF CRC64;

Query Match 99.7%; Score 4051; DB 2; Length 801;
Best Local Similarity 99.8%; Pred. No. 3.8e-240;
Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEALSHVIGIGSPQOLCRLPATYTKLRSHNTSTTIC-SASKMADRLSDFNFTSDSSSS 60
DB 1 MEALSHVIGIGSPQOLCRLPATYTKLRSHNTSTTIC-SASKMADRLSDFNFTSDSSSS 60
QY 61 FATATTTATVLSLPPSIDRPERHVPIDFYQVLAQTHFLTDGIRRAFEARVSKPPQFG 120
DB 61 FATATTTATVLSLPPSIDRPERHVPIDFYQVLAQTHFLTDGIRRAFEARVSKPPQFG 120
QY 121 PSSDALISRQIIOACETLSNPSRREYNEGILDDDEATYTTVPMDKVGALCVLOEG 180
DB 121 PSSDALISRQIIOACETLSNPSRREYNEGILDDDEATYTTVPMDKVGALCVLOEG 180
QY 181 GERTIVLRVGEALIKERLPKSKFQDVVLMALAFDVSRDAMALDPPDFTIGYEFVEBAL 240
DB 181 GERTIVLRVGEALIKERLPKSKFQDVVLMALAFDVSRDAMALDPPDFTIGYEFVEBAL 240
QY 241 KILQEBGASSIAPDLRAQIDETLEETPRVYVLELGLPLGDDVAAKRLNGLSGVNITLWS 300
DB 241 KILQEBGASSIAPDLRAQIDETLEETPRVYVLELGLPLGDDVAAKRLNGLSGVNITLWS 300
QY 301 VGGGGAALVGGLTREKFNENAFPLMTAAEQVDLPVATPSNIPASFEYVEVALVAQA 360
DB 301 VGGGGAALVGGLTREKFNENAFPLMTAAEQVDLPVATPSNIPASFEYVEVALVAQA 360
QY 361 FIGKKPHILOADADKFOOLOAKVMAHEIPAMLYDTRNNMEIDFGELRGICALLIGKVD 420
DB 361 FIGKKPHILOADADKFOOLOAKVMAHEIPAMLYDTRNNMEIDFGELRGICALLIGKVD 420
QY 421 CRMWLGIDSEDSQYRNPAIVEFVLENSNRDNDLPGICKLLETWLAGVPPRFDDTDK 480
DB 421 CRMWLGIDSEDSQYRNPAIVEFVLENSNRDNDLPGICKLLETWLAGVPPRFDDTDK 480
QY 481 KFKIGDYDDPMVLSYERVEVVGSPAAAAAARIGAEHVKASAMQALQKVPSTYTD 540
DB 481 KFKIGDYDDPMVLSYERVEVVGSPAAAAAARIGAEHVKASAMQALQKVPSTYTD 540
QY 541 RNSAPPDQOVETVSVDPVGNNGRDGPGVFIABAVPSENFTNDPAIRAGVSESSVD 600
DB 541 RNSAPPDQOVETVSVDPVGNNGRDGPGVFIABAVPSENFTNDPAIRAGVSESSVD 600
QY 601 ETTVEVSVADMLKEASVYKILAAVAGLISLFSQKYPFLKSSSSFORQDMVSMESDVATI 660
DB 601 ETTVEVSVADMLKEASVYKILAAVAGLISLFSQKYPFLKSSSSFORQDMVSMESDVATI 660
```

```
QY 661 GSVRADSEALPRMDARTAEIVSKQIKSLAFGPDRIEMLEPEVLDGRMLKITDRAA 720
DB 661 GSVRADSEALPRMDARTAEIVSKQIKSLAFGPDRIEMLEPEVLDGRMLKITDRAA 720
QY 721 ETTAQLGVYDYTKLKSVDSTVVSADGTRALVEATLEESACLSDLVHPENNATDVRTYTT 780
DB 721 ETTAQLGVYDYTKLKSVDSTVVSADGTRALVEATLEESACLSDLVHPENNATDVRTYTT 780
QY 781 RYEVFWSKGMKITEGSVLAS 801
DB 781 RYEVFWSKGMKITEGSVLAS 801

RESULT 4
ID 062729 PRELIMINARY; PRT; 760 AA.
AC 062729;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Plastid division protein.
GN Name=P0575F10.2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
CX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa japonica (GA3) genomic DNA, chromosome 2, PAC
clone: P0575F10.2";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP004885; BAD07942.1; -.
DR InterPro: IPR001623; DnaJ_N.
SQ SEQUENCE 760 AA; 84134 MW; 2C44684862795B2F CRC64;

Query Match 43.7%; Score 1775.5; DB 2; Length 760;
Best Local Similarity 47.9%; Pred. No. 2.1e-100;
Matches 390; Conservative 119; Mismatches 213; Indels 93; Gaps 16;

QY 12 SPFQCLRLPATYTKLRSHNTSTTIC-SASKMADRLSDFNFTSDSSSSPATYAT 65
DB 14 APFAFSLPRPRPRRRPRRPPHPSAACAASRMARLPAETHILPTAAPSDPPSAPAPAA 73
QY 66 TTATVLSLPPSIDRPERHVPIDFYQVLAQTHFLTDGIRRAFEARVSKPPQFGSDA 125
DB 74 APSASFPVLPFPDAERSLPIQVDVFKVLAEPHPLDGIIRRAFEARIAKPPQYGYSTDA 133
QY 126 LISRQILOACETLSNPSRREYNEGILDDDEATYTTVPMDKVGALCVLOEGSETEI 185
DB 134 LVGRQMLQAHDTLNMQNSRTQYDRALSENREBALTMWDAMD-----SAGEALA 184
QY 186 VLARGBALIKERLPKSKFQDVVLMALAFDVSRDAMALDPPDFTIGYEFVEBALKILQ 245
DB 185 VLVTGEGLLDPRPKRKQDVVLMALAYVDSRDMAAPSPDVIGCCFVLERLKLLOE 244
QY 246 EGASGLAPDLRAQIDETLEETPRVYVLELGLPLGDDVAAKRLNGLSGVNITLWSVGGG 305
DB 245 DGASNLAPDLISQIDETLEETPRCVLELPLPDTHEHKRROGLGAGRIILWSVGRG 304
QY 306 ASALVGLTREKFNENAFPLMTAAEQVDLPVATPSNIPASFEYVEVALVAQAFIGKK 365
DB 305 IATVGGFSRAAFENAFPLMTSIEQMDFPSKTPNSIPPEWFEIYVVALAHVAQAIIKSR 364
QY 366 PHLQADADKFOOLOAKVMAHEIPAMLYDTRNNMEIDFGELRGICALLIGKVDCEKML 425
DB 365 PPIIMADDLPEQLOKKNIGS----HYAYDN----EVDLALERFCCGLIVDVSKCMM 416
QY 426 GLDSEDSQYRNPAIVEFVLENSNRDNDLPGICKLLETWLAGVPPRFDDTDKFKL 484
DB 417 GIDVSESPYDPKILFEIYVNSISEENDLPGICKLLETWLIPEVPPRSRDTGMOFRL 476
```

Qy	485	GDYDDPWLSTLKEVEVYQSSPLAAATMARIAG-----HYKSAOALQKPEP-SR	537
Db	477	GDYDDPEVLTSLERMEGGASHLAAAAIADGAGATATALGTVKSNAIQAFKPEPLE	536
Qy	538	YTRNSAEPKQVQETVFSVDPVGNVGRDGPQVFIAEVRPSENEPTNDYAIRAGVSES	597
		:	:
Db	537	QUDRSAMENT-----KQG-PEGY-----ENFDQ-----	559
		:	:
Qy	598	SVDETTEVMSVADMLKEASVKILLAGVALIGLISLFQKTFKSSSSSFQRDQWVSMSPD-	656
		:	:
Db	560	-----ENAPHRDNRNALKTIISGALPALLAVIGAKI-----LPRRPPLSARSEH	605
		:	:
Qy	657	--VATIGSVRADDSAL-----PRMDARTLENVSKKQKIKSLAFGPDHRIEMLPEVL	707
		:	:
Db	606	GSVAVAANSVDSITDDPALDEDPVHIIPMDAKTLAEIDIVRKQSIISKALGPEHSAVSLQEVY	665
		:	:
Qy	708	DGRMLKITDRAAEFNAQLGVVDYITLKLSVDSVYVSADGTALVATITLESACLSLDYH	767
		:	:
Db	666	DGNMLKVTDRAAELERHGMFEWETYSIDVTIDSIITSLDGRRAVYATIDBAGQLDVTY	725
		:	:
Qy	768	PENNAATDRTVYTRTYREVFMSK-SGMKITGSGYLAS	801
		:	:
Db	726	PRNDSYDTKITTYREMAFSKLGKGMKITLGVALKS	760
		:	:

ID	Q7PC78	PRELIMINARY;	PRT;	760 AA.
AC	Q7PC78;			
DT	01-MAR-2004 (TREMBLrel. 26, Created)			
DT	01-MAR-2004 (TREMBLrel. 26, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Plastid division protein.			
GN	Name=Arc6;			
OS	Oryza sativa (indica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzae; Oryza.			
OX	NCBI_TaxID=39946;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RX	MEDLINE=22779057; PubMed=12897262.			
RA	Velha S., Froehlich J.E., Kosharova O., Pyke K.A., Van Erp H.,			
RA	Osteeroung K.W.;			
RT	"ARC6 is a J-domain plastid division protein and an evolutionary			
RT	descendant of the cyanobacterial cell division protein Fz2.";			
RL	Plant Cell 15:1918-1933(2003).			
CC	-I- MISCELLANEOUS: The sequence shown here is derived from an			
CC	EMBL/genbank/DDBJ child party annotation (TPA) entry.			
DR	EMBL; BK000929; DAA01472.1; ..			
DR	Gramene; Q7PC78; ..			
DR	InterPro; IPR001623; DNAS_N.			
DR	SEQUENCE 760 AA; 84134 MW; 2C4684862795B2F CRC64;			
QY	Query Match	43.7%;	Score 1775.5;	DB 2; Length 760;
	Best local similarity	47.9%;	Pred. No.2.1e-100;	
	Matches 390;	Conservative 119;	Mismatches 213;	Indels 93; Gaps 16
QY	12 SPFLCRLPPTTKLRSHNTSTTC-SASKWADRLSLDFNF-----TSDSSSSSFATAT 65			
Db	14 APFAFSLPRPRPRRRRRPPHPSAACRAASRWAEKLFADFHLLPTPAASDPSPAPADAA 73			
QY	66 TTAIVLSPPSIDREHVRPIPIPIYQVIGAGTHTTGIGIRAPARVSRKSPGFSNDA 125			
Db	74 APSASFPVPLPDDAERSLPQVDPDYKVLGABPHFLVSGIRAFARIAKPPQGYSTDA 133			
QY	126 LISRRQLQAACETLSNPSRSREYNEGLLDBEATVITDVPMDKVPGALCVLOEGGETEI 185			
Db	134 LVGRKQMLQIHDTLTMNQNSRTOYDRALSENREKALTMDIAMDK-----EAGELAA 184			
QY	186 VLRGEALLKRLPKSFQDQVYLVALLAFLDVSRDAMALDDPPDFTTGVEVFEALKLQIE 245			

Db	189	VLTGSGQLLDBRPPKRFKQDVVLAMALAYVDLSRDNAMAASPPDVIIGCEVLERALXULQJ	244
Qy	246	EGASSIAPDLRQIDTLEETLETPRYVLELGLPLGDDYAAKRLNGLSGVRIILMSVGGG	305
Db	245	DGASNNAPRLLSQIDTLEETLETPRCVLELSPIDTLEHHKROGLOGARNIILMSVORG	304
Qy	306	ASALVGLTRERKPMNEAFLPMTAAEQVDLPVATSNTPAESFEFEYEVVALVVAQAFIGKK	365
Db	305	IATVGGGFSREAPMNEAFLPMTAISQMPFSKTPNSLPPEFEFEYVALVVAQAIISKR	364
Qy	366	PHLLQADAKFOOLQAKVAMEIEMMDYDRNNWEIDFGELGLCALLIGYDECRMWL	425
Db	365	PGYIMADDLFPLQIQFNIGS----HYAYDN-----ENDLALERAFCSLIVGVSCKRMWL	416
Qy	426	GLDSEDSQIRNPAIVFPLVLENN-RDDNDOLPLGLCYLLETPLAGVPPRRDTRKOKKFKL	484
Db	417	GIDNESSPRDRKILFEIYVNTSSISEENDLPLGCKLETLIEVPPRRDTRGQMFRL	476
Qy	485	GDYDDPMTLYTERVEVVOGSPLLAAATMRLGAE-----HYKASAMQALQKFP-SR	537
Db	477	GDYDDPPEVLTSLTERKEGGASHLAAAAIKLGAQAPALGLTVKSNALIQAFNKPVLIE	536
Qy	538	YTDNRASBPKDVOETVFSVDPVGNVNGRDEGPVFLAEVAPSENPFETNDYAIRAGVSES	597
Db	537	QIDRSAMENT-----KDG-RGGYL-----ENFPQ-----	559
Qy	598	SVDETTEVMSVADMKEASVYKILAGVAILGLISLPSQKYPFKSSSSSQORDMVSMSMD-	656
Db	560	-----ENAPHDSRNNALKTISAGALFALLAVALGATY-----LPRKRPPLSARSEH	605
Qy	657	--VATIGSVRADDSLEAL-----PRMDATAEINIVSMQKIKISLAGPDRILEMPEVL	707
Db	606	GSAYAVANSVDSITDDPDLDBEDPVHITRPMKAKLAEIDYKRMQSIKSKALGPBHSVASLOEVL	665
Qy	708	DGRMLKIWTDRAAETAGLGVYDYVTLKLSDVSQTVVSSADGTRALVEATLEESACLSULVH	767
Db	666	DGNMLKVMWTDRAAETIERHGMFEWETLSDVITDSITISLIDGRRAVTEATIDEAGQLDVTE	725
Qy	768	PENNAATVATYTTTRYEFPMK-SGKIKTEGVLAS	801
Db	726	PRANDSYDTRKYYTRRYEMAFSKLGGKMLTEGAVLVS	760

RESULT 6			
Q7XAS1			
ID	Q7XAS1	PRELIMINARY;	PRT; 324 AA.
AC	Q7XAS1;		
DT	01-OCT-2003 (TREMBLrel. 25, Created)		
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)		
DT	01-MAR-2004 (TREMBLrel. 25, Last annotation update)		
DE	Truncated division protein.		
CN	Name=arc6;		
CS	Arbidiopsis thaliana (Mouse-ear cress).		
CC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
CC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
CC	eucotsids II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_TaxID=3702;		
FN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22779057; PubMed=12897262;		
RA	Viltha S., Froehlich J.E., Kosharova O., Pyke K.A., Van Erp H.,		
RA	OsterYoung K.W.;		
RT	"ARC6 is a J-domain plastid division protein and an evolutionary		
RT	descendant of the cyanobacterial cell division protein Ftn2."		
RL	Plant Cell 15:1918-1933 (2003).		
DR	EMBL; AY221467; AAQ18644.1; -.		
DR	InterPro; IPR001623; DnaJ N.		
SQ	SEQUENCE 324 AA; 35468 MW; OCC5C0CC469A6325 CRC64;		
Query Match 40.5%; Score 1646; DB 2; Length 324;			
Best Local Similarity 100.0%; Freq. No. 5,4e-93;			
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

```

QY 1 MEALSHVIGISLPOLGRLPATTKLRSHNTSTTICGASAKMADRLSDENFTSDSSSS 60
Db 1 MEALSHVIGISLPOLGRLPATTKLRSHNTSTTICGASAKMADRLSDENFTSDSSSS 60
QY 61 PATATTTATVSLPSPIDRPERHVPIDFYQVIGAQTHFTLTDGIRRAFEARVSKPQFG 120
Db 61 PATATTTATVSLPSPIDRPERHVPIDFYQVIGAQTHFTLTDGIRRAFEARVSKPQFG 120
QY 121 FSDALISRRQIIQAACETTSPRSRRYNEGILDDDEATYITVPMDKVGALCVLOEG 180
Db 121 FSDALISRRQIIQAACETTSPRSRRYNEGILDDDEATYITVPMDKVGALCVLOEG 180
QY 121 FSDALISRRQIIQAACETTSPRSRRYNEGILDDDEATYITVPMDKVGALCVLOEG 180
Db 121 FSDALISRRQIIQAACETTSPRSRRYNEGILDDDEATYITVPMDKVGALCVLOEG 180
QY 181 GETEIVLRVGEALKEKRLPKSPKODVVLVMAALFLDVSADAMALDPDFITGEVEBAL 240
Db 181 GETEIVLRVGEALKEKRLPKSPKODVVLVMAALFLDVSADAMALDPDFITGEVEBAL 240
QY 241 KLIQEBGASSIAPDLRAQIDETLEITPRYVLELGLPLGDDVAAKUNGLSGVRNITLMS 300
Db 241 KLIQEBGASSIAPDLRAQIDETLEITPRYVLELGLPLGDDVAAKUNGLSGVRNITLMS 300
QY 301 VGGGASALVGLTRKEMNEAFL 324
Db 301 VGGGASALVGLTRKEMNEAFL 324

```

RESULT 7

```

ID 08YTL0 PRELIMINARY; PRT; 798 AA.
AC 08YTL0;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE A112107 protein.
GN OrderedLocustNames=a112107;
OS Anabaena sp. (Strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Matanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
RM EMBL; AP003590; BAB74406.1; -.
DR PIR; AD2144; AD2144.
DR HSP; P08622; IBOZ.
DR Pfam; PF00226; DnaU.1.
DR SMART; SM00271; DnaU.1.
DR PROSITE; PS50076; DnaU_2.1.
KM Complete proteome.
SQ SSOURCE 798 AA; 90124 MM; 3C57B66C21722AA CRC64;

```

Query Match 12.24; Score 497; DB 2; Length 798;

Best Local Similarity 24.64; Pred. No. 7.9e-22;

Matches 207; Conservative 139; Mismatches 301; Indels 196; Gaps 33;

```

QY 84 VPIPIDPVYVGAQTHFTLTDGIRAPARVSKPQFGSDALISRRQIIQAACETTSPN 143
Db 11 VPIPIDPVYVGAQTHFTLTDGIRAPARVSKPQFGSDALISRRQIIQAACETTSPN 143
QY 144 RSRREY-----NEGILDDDEATYITVPMDKVGALCV 176
Db 71 KESSYQVLAIAHAYPDNATTKVAVENGDNSNGHFDVQSLSI--EVSSEELIGALLI 128
QY 177 LQSGGEIYLRVGEAL-----KERLPKSF-----KQDVVLMALAPLDVSR 219
Db 129 LQSGGEIYLRVGEAL-----KERLPKSF-----KQDVVLMALAPLDVSR 219

```

```

QY 220 D-----AMALDPDEFITGEVEBALKLIQEBGASSIAPDLRAQIDETLEITPR 269
Db 189 EQMGQGHYENALSL-----TGE-----VLFSG-----IFSVQABIQADLYKLRPY 233
QY 270 YLELLPLGDDVAAKUNGLSGVRNITLMSVGG--GASALVGLTBE--KFNENAFR 325
Db 234 RILELLALP--OEKTIIRHOGDLLOSILDRGIDGTGNDOSGLNIDDFLRFIOQRH 291
QY 326 MTAEOVDLVAPRPSNIPARSEVEVALVVAQAFIGKKPHILODADKQFOQIQAKRM 385
Db 292 LTVABQKLPDG--ESKPS--AVATYLAIVASIAQFQROPALIRHAKQILMRISKQ-- 347
QY 386 AMEIPAMLYDRNNWEIDFGLERGLCALIGKVDCEMVLGLDSEDSQYRNPAIVEVLE 445
Db 348 -----DVHESGLCALLGGTEEARVLESGE--YEALALI-----R 383
QY 446 NSNRDNDLPLGLCKLETLVAGVPPRFDYDKKFKLGDYDDPNVLSYLEV----- 500
Db 384 EKSQSDPLPLGLCTVAEQWLQNEVFPHPRDLRSQASLKDQYFANQOVQALTEALPDAE 443
QY 501 -----EVQGSPLAAATMAR--IGAEH--VRASAMQALQKVPSS--RYTD 540
Db 444 TTNEWAVINQSPQPRGNSYSGTVPAKRPVGKAPRGEASTRPVQKSHPSVNRQFH 503
QY 541 RNSAPKQVOETVFSVDPVGNVGRDGE-----PGVFLAAVRPSSENFETND 587
Db 504 QNRTPDELPETSHHRPSSNFTTARENISTDYATDNYPELPVRSAPR----- 555
QY 588 YAIRAGVS--ESSVDFTV-----ENSVADMLEKASVKILAA 622
Db 556 --VQPGVSGYQSTPPRQTPKRRRKKPQAVNRGHSIHQOROPSPSTLRKTLIVL 613
QY 623 GVAIGLSLF-----SQKX-FLKS-----SSSFQRDMVSMESDVATIGSVRADSEALPR 673
Db 614 G-SIGGILLFWLVTTFGWLKNVFPFAPSLQSGQLSIQSPLEIPDKNAQIQSEPV 672
QY 674 MDARTAEIVSKQKIKSLAFGPHRIEMLEPVLDGRMLKMTDRAETAQGLVVDYTL 733
Db 673 LTERTAKIEMTLATGASALGAHEKIESLEILITGSLAQWRILALQDADRNRHREYS- 731
QY 734 LKSVDSVTVS-ADGTRALVEATLESACISDLVH--PENNAVDVRYTTRYEVFWSKG 790
Db 732 HSYVDSISKSIDIDPNASVGATVRE--LTQFYENGQKKSDEP--LRVRYELIRODDI 787
QY 791 WKI 793
Db 788 WRI 790

```

RESULT 8

```

ID 07NEP3 PRELIMINARY; PRT; 626 AA.
AC 07NEP3;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE G113481 protein.
GN OrderedLocustNames=g113481;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]
RP SEQUENCE FROM N. A.
RX STRAIN=PCC 7421;
RM MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Matanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids."
RL DNA Res. 10:137-145(2003).
RM EMBL; AP006580; BAC91422.1; -.

```

DR InterPro; IPR001623; DnaJ_N.
 DR InterPro; IPR008941; TPR-like.
 DR Pfam; PF00226; DnaJ; 1.
 KW Complete proteome.
 SQ SEQUENCE 626 AA; 69297 MW; 372BAB258B84629 CRC64;

Query Match 11.0%; Score 448.5; DB 2; Length 626;
 Best Local Similarity 25.8%; Pred. No. 5.3e-19;
 Matches 192; Conservative 112; Mismatches 282; Indels 159; Gaps 30;

QY 86 IPIDFYQVLAQTHFLTDGIRRAFEARVSKPPGPGSDDALISRRQLQAACETLSNPRS 145
 DB 9 LPLSYQIILCVGPOCTYQVEPAADLAQAPREFSAAVSAHEHMLREKCTALGPVR 68
 QY 146 RREYN-EG---LLDDEATVITDVPMDKVPALCVLOEGGETEIVLRVGEALIKERLPK 200
 DB 69 RELYHREGKGLLSDSHAAV-----GLFLVELGEYQILIRQREALAVANQP- 117
 QY 201 SFKQDVVLYNALAFLDVSRDAMALDPPDFTGYEFVEBALKLOEGGASSIAPDLRAQID 260
 DB 118 ----DTRLVLAIAHQALAQEAYR--QGNLTLAHLEVEALEILR--GDDCIKP-VQDELQ 168
 QY 261 ETLEETPRVYLIELGLPGDDYAKRLNGLSGVNLT---MSVGGGASALVGLTRE- 316
 DB 169 TLKRWPRERILQILA-GAADPPSPQOEGWALLAALLAEREIEGIDGNDQ--SGLSREE 225
 QY 317 --KFMNEAFPRMTAAEQVDLF---VATPSNIPAESFEVYEVALLVAQAFIGKKPHLLQD 371
 DB 226 FVGFLOVLRRLRYABEQBELFERAARPS--PAHQ---YLAQAQALACGFTEGSPQVR 280
 QY 372 ADKQFQQLQAQKVMAMEIPAMLVDTNNWEIDFGLERGLCALLIGKYDECRMILGLDSED 431
 DB 281 ARGHLIKLVORQ-----DVLLELAVCALLIGQVEAOKNIERSABE 321
 QY 432 SQKRNPAIVFVLENSNRDNDLPGCLKLETLVAGVPRPRTDOKKPKGDYDDP 491
 DB 322 Q-----AVDYI--KQLSQDSPDLPGLCRYTDLAEVFPGFDRRSCTYTLKAYFAHP 374
 QY 492 MVLTYLERVAVVQSGPLAAATVMARIGAEHYKASAMQALQKFPSPRTDNRSAEPKQVE 551
 DB 375 EVAAPLD-----DPPAPAPAPERPAPVIASGAS-----CTIERGL-- 413
 QY 552 TVESVDVGNVNGRGEPPGVFIABAVRPSSENFETNDYAIRAGVSESSVDETTVMESVAD 611
 DB 414 ----LTFVG-----AEARRR-----ANL 428
 QY 612 LKASVYKILAAVAILGIS--LFSQKFLKSSSPQKQWVSMESDVATIGSVRADDS 668
 DB 429 LSPQASVAAALVAIFLLSGGWLTTOROSRPEPPAPAP--VAPPVQAPAPIAAPAPQ 486
 QY 669 EALPR-----MDARTAEINIVSKQIKSLAFGPDHRIEMLPVLDGRMLKIWTDR 718
 DB 487 AAAPFLTAADNDQPTDAQIATA--MLKNWQTKAKQALQEPHERTAQMOMLTGSSPQRVQOK 545
 QY 719 AAEATQGLVYDYTLKLSDSVTSVAD--GTRALVEATLEESACL--SDLVHPENNATD 774
 DB 546 VEQSRQAGEYWKSESLKDLKIEQV--ADRRPDVAVAQVTEVAVNLTYDQLAPRSRSD- 601
 QY 775 VRTYTRRYEVFMSGKMITESVLT 799
 DB 602 -RPPRVKISLVKAPAGMRIEEMKVL 625

RESULT 9
 Q93AKO PRELIMINARY; PRT; 631 AA.
 AC Q93AKO;
 DT 01-DEC-2001 (Tremblere). 19, Created)
 DT 01-DEC-2001 (Tremblere). 19, Last sequence update)
 DT 01-MAR-2004 (Tremblere). 26, Last annotation update)
 DE Cell division protein Ftn2.
 OS *Synechococcus* sp. (strain PCC 7942) (Anacyetis nidulans R2).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.

OX NCBI_TaxID=1140;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PCC 7942;
 RX MEDLINE=2206501; PubMed=12218043;
 RX DOI=10.1128/JB.184.19.5524-5528.2002;
 RA Kosharova O.A., Molk C.P.;
 RT "A novel gene that bears a DnaJ motif influences cyanobacterial cell
 division."
 RT J. Bacteriol. 184:5524-5528(2002).
 DR EMBL: AF421196; AAL16071.1;
 DR GO:0000910; P:cytokinesis; IEA.
 DR InterPro; IPR001623; DnaJ_N.
 DR Pfam; PF00226; DnaJ; 1.
 DR SMART; SM00271; DnaJ; 1.
 DR PROSITE; PS50076; DnaJ_2; 1.
 KW Cell division.
 SQ SEQUENCE 631 AA; 70693 MW; 6E46142B48DEB2 CRC64;

Query Match 10.9%; Score 442; DB 2; Length 631;
 Best Local Similarity 24.5%; Pred. No. 1.3e-18;
 Matches 189; Conservative 118; Mismatches 262; Indels 204; Gaps 30;

QY 86 IPIDFYQVLAQTHFLTDGIRRAFEARVSKPPGPGSDDALISRRQLQAACETLSNPRS 145
 DB 3 IPIDYRILICVQVQASADKLAEYSRDLNQPSEHFESELAQRRQLLEALIELSDPEQ 62
 QY 146 RREYN-EG---GLDDEATVITDVPMDKVPALCVLOEGGETEIVLRVGEALL----- 194
 DB 63 RDRYDRRFGQGLEAIEPSIELED--WQRI-GALLILIEIGEYDRYSQALAEILLPVDAS 119
 QY 195 KERLPKSFQDVVLYNALAFLDVSRDAMALDPPDFTGYEFVEBALKLOEGCA----- 248
 DB 120 AEVRDQFARGDIALAIALSQOSIGRECRQ-----QGLVEQAQHQHGRS 162
 QY 249 -SSLA-----PDLRAQIDETLEETPRVYLIELGLPGDDYAKRLNGLSGVNLT---W 299
 DB 163 QSALADHQRPPELSRTLHQOQGLRPRYILIERLAQPLTD--SDRQGLILLAMDRLQ 220
 QY 300 SVGGGASALVGLTREKFMNEAFLR-----MTAEQVDLF---VATPSNIPAESFEVY 351
 DB 221 GIEGPEDDG--SGLTIDNFI--MFLQOIRGYLTLAQQLFSEARARPS--PAASF---F 271
 QY 352 VALAVQAQFIKPKPHLLQDADKQFQQLQAQKVMAMEIPAMLVDTNNWEIDFGLERGLC 411
 DB 272 ACYTLIARGFCDQPSLIHRSALLHELKS-----RMDVHIQQAIA 312
 QY 412 ALLIGKYDECRMILGLDSEB---SQYRNPAIVFVLENSNRDNDLPGCLKLETLWIA 467
 DB 313 SLILGQPEEAELL--VOSQDEETLSQIRALAQGEAL-----IVGLCFETWIA 360
 QY 468 GVFPFRDRTDKKFKLGDYDDPMVLYLERVAVVQSGPLAAATVMARIGAEHYKASAM 527
 DB 361 TKYFPDRDKETATLQPIFDPPDQVYIDALVEL----- 396
 QY 528 QALQKVPKRYTDRNSAEPKQVQTVFSYDP-----VGNVNGRGEPPGVFIABAVRPS 580
 DB 397 -----PSDLMPTPLPVPLEVRSSLLAKELPTPATRP-----VAPP 432
 QY 581 ENFETNDYAIRAGVSESSVDETTVMESVADMLKEAIVKILAAVAILGISLFSQKFLKS 640
 DB 433 PRRRRDRSERPARTKRPLPMIGLV-----VVVLGGGTGV-----NANRS 475
 QY 641 SSGFRQKQWVSMESDVATIGSVRADSEALP-----RMDARTAEINIVSKQIKS 691
 DB 476 RSN-----STPPTPRPVQGLPBAVPRAPAPVVALDRQAQALVVLQNMIAAKA 524
 QY 692 LAFGPDHRIEMLPVLDGRMLKIWTDAETA--QGLVYDYTLKLSDSVTSVADGTR 749
 DB 525 AALGPGYDRRLATVLTGAVLTQWQGFSSQAANTQLTSGDH---KLYVDSVQLSDGDR 581
 QY 750 ALVBAILEBACSLDVHPENNATDVRT--YTRRYEVFMSGKMITESVLT 799
 DB 750 ALVBAILEBACSLDVHPENNATDVRT--YTRRYEVFMSGKMITESVLT 799

Db 582 AVQAKVDE----VEQVYRGDQLETRRDGLVIRYQLVRENNIMKIASISLV 630

RESULT 10

Q8VY16 PRELIMINARY; PRT; 819 AA.

AC Q8VY16; 01-MAR-2002 (Tremblrel. 20, Created)

DT 01-MAR-2002 (Tremblrel. 20, last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, last annotation update)

DE Hypothetical protein Ac3g19180.

GN Name=Ac3g19180;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI_TaxID=3702;

OX [1]

RA SEQUENCE FROM N.A.

RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,

RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,

RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,

RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,

RA Seki M., Shimizu P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,

RA Theologis A.

RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY074283; AAL6980.1; -.

KW Hypothetical protein.

SO SEQUENCE 819 AA; 90750 MW; 6720083F4864657E CRC64;

Query Match 10.6%; Score 431; DB 2; Length 819;

Best Local Similarity 23.5%; Pred. No. 9.4e-18; Indels 224; Gaps 32;

Matches 200; Conservative 116; Mismatches 311;

QY 56 SSSSFATATTATVTLVSLPPSIDRPERHVPIDIFYOVLAGNQTFTDGIKRAFEARVSK 115

DB 81 SRTSSLAASST-----ILPVCYQLIGSEQADEVKSINLKT 124

QY 116 PPGGFSDDLISRROQLAACETLSNPSRREYNEGLDDEATVTTDVPWKVQALC 175

DB 125 DAEQGYMEAAARQDLMDVDRDL--LPESEYAGNKEKIKAPSLRIPFWMLPGALC 181

QY 176 VQEGGTEIVLRGELTKERLPKSPQVVMALAFDVSADAMALDPDPTGYEF 235

DB 182 LLDVGQEKVLVDIGRALRNLDSKPYTHIDIFLSMALBCAIAAAEVNVS--QGFEA 239

QY 236 VREALKLQEE-GASLAPDLRAQIDETLEBITPRVYELTGLPLGDVYAKRLNGSGV 294

DB 240 LAAQSTLKSRYTLGKLA--LTQIESLBEGLAPCTLDLGLPRTENMERKGAIAL 297

QY 295 RNILMSVGGGASALVGLTRK-----FNNEAFLEMTAAEOYDLF-----VATP 339

DB 298 RELLRQ-----GLAVEASCQIQDWPCFLSQMISRLLAIEYIDLFWDDLATRK 346

QY 340 SNTPASEFE-----YEVALLVAQAFIGKPHLLQADQFQLOQAKMAEIP 390

DB 347 NKGLSESHNORVVIDPNCFTVNLGHIAVFGS-----KONETINAKTICCLI 396

QY 391 AMLYDTNNNEIDPGLERGLCALLIGKVDCEKMWIAGL-----DSEDSQYRNPAIYEVLEN 446

DB 397 A-----SEGVDLKFEAPCSFLKQSGSEBALEKQLESNDSAVRNS-----ILGK 444

QY 447 SNRDDNDLPGCLKLTWTLAGVPPRPRTK-----DKPKLGDYVDPMVL 494

DB 445 ESRSTS-----ATPSLAWMLMESVLANFPDTRGCSPLANFFRAEKYKPKKMGKSPSII 499

QY 495 SY-----LRRVRYOOSPLAAATMARIGAENRYKASGMOLQVPPSRRTYDRKSAPKO 548

DB 500 NHTKQRPSTTQFVNS-----OHL-----YTAVEQLTPTD 531

QY 549 VQETVFPVDVGNVNGVGDGEPGFIAAVRPSSENFETNDYAIRAGVESSV-DETVEMS 607

Db 532 LQSEFVVSAAK--NNDE-----TSASMPYQVKEN-----LGVAKNKIMDEM-----LS 571

QY 608 VADMLKASVKKILLAGVAIGLISLFSQKY-----

Db 572 QSSLIGRVSVAL-----LCCTVFFSLKLSGINSGLQSPISVSNRPHSEDSFLMTE 626

QY 637 ---FLKSSSFORKDVS-----MESDVATIGSVRADSEALPR 673

Db 627 SGNFRKMLDSVNNNGIYGNIKVLIDMLKMGHGEHPALYLKSSQASATSLSHSASELHKR 686

QY 674 -MDARTENIVSKKQKISLAFGPDRILEMPEYLDGRMLKITDRALFTRQLGLV-DY 731

Db 687 PMDTEBAEELVROENYKAAALGPTHQVYSLEVLDESMTVQW-QTLAQTAARAKSCWRP 745

QY 732 TLKLKLSVDSVTSAD--GTRALVEATLESASACSDLVHPENNATDVRTTYTRVEFMSK 788

Db 746 VLHLLEVLQAHIFEDGJAGNAAKIALLEBAELVDSQPK-NKTYSTYKIRIYLKQK 804

QY 789 SG-WKITEGSV 798

Db 805 DGLMKFCQSDI 815

RESULT 11

Q5559 PRELIMINARY; PRT; 714 AA.

AC Q5559; 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, last sequence update)

DT 01-JUN-2003 (Tremblrel. 24, last annotation update)

DE S110169 protein.

GN OrderedlocusNames=s110169;

OS Synechocystis sp. (strain PCC 6803).

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OC NCBI_TaxID=1148;

OX [1]

RA SEQUENCE FROM N.A.

RC STRAIN=PCC6803;

RX MEDLINE=97061201; PubMed=8905231;

RA Kaneo T., Sato S., Kocant H., Tanaka A., Asamizu E., Nakamura Y.,

RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,

RA Hoshouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,

RA Shimizu S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,

RA Tabata S.

RT "Sequence analysis of the genome of the unicellular cyanobacterium

RT Synechocystis sp. strain PCC6803. II. Sequence determination of the

RT entire genome and assignment of potential protein-coding regions."

RL DNA Res. 3:109-136(1996).

DR EMBL; D63999; BAA10060.1; -.

DR PIR; S76082; S76082.

DR HSP; P08622; 18QZ.

DR InterPro; IPR001623; DnaJ_N.

DR Pfam; PF00226; DnaJ_1.

DR SMART; SM00271; DnaJ_1.

DR PROSITE; PSS0076; DnaJ_2; 1.

KW Complete proteome.

SO SEQUENCE 714 AA; 79423 MW; 51B52C16F405ED3E CRC64;

Query Match 10.6%; Score 429; DB 2; Length 714;

Best Local Similarity 25.1%; Pred. No. 1e-17;

Matches 194; Conservative 124; Mismatches 331; Indels 124; Gaps 30;

QY 86 IPIDFYVLAGNTHFLTDGIRRAFEARVSRPPGFSDDLISRROQLAACETLSNPS 145

DB 3 IPIDFYRILIDIPPOSGETIEQAYODRLQLPRREFSDAAVTYLNQLLAAYETLRDEK 62

QY 146 RREYNE--GLDDE--EATVITDVPMDKVP-----GALCVLQEGGETEIVLRGEALLKE 196

DB 63 RQAYDQEWGAMBALEALPITPTELECSPPQEGIGALLILDGEGELVVKYGEPLIHD 122

QY 197 RLPSK--FKQDVVLMALAFDVSADAMALDPDPTGYEFVEEA-LKLQEGASSLAP 253

DB 123 PNPAGGLPDYLLSVILAHMELSRERWQOO-----YFAATASIKALARLQODNDP 176

ID	PRELIMINARY	PRT	841 AA.
QY	254	DLPAQIDETLEITPRRYVLLGLGP-LGDQYAAKXNGISGRNITMSVGG-GGASALVVG	311
Db	177	ALAEATIRELIRKRYRIITELLAKGSGEE---QHQGGALLQANVDRGEGEGEDYS	233
QY	312	GLTRE--KPNNEAFRLMTAAEQVDLFVATPNSNIPAESFEYVALVAQAIFKKCHL	368
Db	234	GLGNDPFLKFTIHLCHLTVAEQNALFLP--PESQRPISVASTAVASHISMAEIVXKODPWA	291
QY	369	LODADKQFOQLQAQAVMAEIPAMLYDRNMWEIDFGLERGCALLIGKVDCEBWMGLD	428
Db	292	IWEAKSLIQLENCQ-----DLAEKVICELLQCTEV--VLAIID	330
QY	429	SEDSQYRNPAIVEFYLENSNRDNDLRLCLLETWLAGVFPFRDTRKQKFKLGDY	488
Db	331	QGD-----PKIVA-GLSEKLTATGDEPLTAFTYFTQEWLEBEIVFYRDLSPETLSPKAYF	384
QY	489	DDPMVLTSLREVEVQ-----GSP-LAAATMARIGAHNKASAM-QALQKPPRSYT	539
Db	385	NNPSYQQLQLEQLEPDSFTTDSFAPALSLTSTSETPMVHSSALPDRPLSTVPS---	441
QY	540	DRNSAPEDVQETVFSYDPPVGNVYRDRGEPGVFIEMAVRPSNEFTNDYAIRA-----	592
Db	442	-RRGSPRRASNDVDP-----SADNSSLGAYT-TLSPALYHTSLGNTGIGDSTS	491
QY	593	-GVSESSVDEFTTVE-MSVADMKKEASVK-----IIAAGVALGILSLFSOK-----	635
Db	492	NGFSSNSPABESTSKKSKSPRRRKRRTIKIPVRFGILCLAGVGGATLAINRTGDPILGG	551
QY	636	-----YFLSSSSFFQRKDMYSMESVATIGSVRADSEALPMDARTANIYSKKMQ	688
Db	552	LLEBDPLVFL-----DQSEETIPREBATSRLILSQ----PFPNQVQGMVQVGLD	598
QY	689	IKSLAFGPDHRIEMIPVLDGRLMKITWDRAEFTQRLGELYDTYLLKLSVDSVTYS-ADG	747
Db	599	SKKLAFGQNYVVGALQSVTLAEPILLAQQRGR-AQRDQAKVNYQYENKQILAIAYQVNPDP	657
QY	748	TRALVEATLEESACLSLDVHPENNATDVR-TYTTREYEFWSKSGKITEGSVL	799
Db	658	NRAIVTAAVEEISQPTFLGNQQAQKSAKRDLLTVAYQLVRRHQGVKILQIQIYV	710

Query Match	10.5%	Score 425	DB 2	Length 841
Best Local Similarly	23.0%	Pred. No. 2.3e-17		
Matches 200	Conservative 122	Mismatches 308	Indels 240	Gaps 35

Qy	56	SSSSSPATTTATVLSPPSIDRBRHVPPIPIPIFYQLGAQTHFLDQIRRAFEAVSK	115
Db	81	SRSSLSASTSI-----IELPYCQILGVSEQAEDVYVYNLKT	124
Qy	116	PPQFGSDALISRQIIOACETYSNRSRRENEGILDEEXTVTVTDVPMKVPQALC	175
Db	125	DAEEGYEMAAARDDLMDVADKLI--LPSEYVAGNIKKEIAKSPILRIPMWLPALC	181
Qy	176	VLOEGETEIVLRVGBALLKERLPSPKQDVVLWVALFLDVSDDAALPDPDTYEF	235
Db	182	LLOEVQCKVLIDIGRALRNIDSKPYTHDIFLSMALAECAIQAALFVVKVS--QGFEA	239
Qy	236	VEBALKILQEE-GASSLAPDLPAQIDETLBEITPRVYLELGLPLGDYAAKRLNGSV	294
Db	240	LAQAQFSLKSKYTLGKLA--LLTQEESELEIAPCTLDLGLPRTPEMERRGAIYAL	297
Qy	295	RNLTMSVGGGASALVGLTREK-----PMNEAFIEMTAEOVDLF-----VATP	339
Db	298	RELIRKQ-----GLSVEASCOIDMPCFSQASIRLAEIYIDLPLMDLATTCK	346
Qy	340	SNIPASFE-----VVEVALVALQAETIGKKPHLLPADKQFOLOQAKYAMEIP	390
Db	347	NKKSLESHQNRVIDPNCFYWLIGHIAVGSG-----KQNTINKAKTIGCCL	396
Qy	391	AMLYOTRNWEIDFGLERGLCALLIGKYDEC--RMWLGJDS-----DSQY	434
Db	397	A-----SEGVDLKFEEAFCSFLLKQLSATGPTCWAOGSEABALEKLNQLESNSDAV	449
Qy	435	RNDALVEFLJENSDDNDLPGCLKLETWLAGVFPFRPDKT-----DKKF	482
Db	450	RMS-----ILGESRSTS-----ATPSLEAMLMESVLANFPDTRGCSPLIANFPRAEKY	499
Qy	483	KLGDDYDDPMVLSY-----LERVEVQGSPLAALAATMARIAGEHVKASAMQLOKYFPS	536
Db	500	PEWKMGSPSIMMHNKTQRPILSTTOFVNS-----QHL-----	532
Qy	537	RYNDRMSAPKVOETVPSVDPVGNNGVRODEPGVFIAEARPSBPENFETDVAIRAGVSE	596
Db	533	-YTAVEQLTPTDLOSPPVSAK---NNDE-----TSASMBPVOLKRN-----LGYHK	574
Qy	597	SSV-DETTYEMSADMLKEASYKIL-----AGVALIGLISLF-----SQ	634
Db	575	NKIWDEM--LSQSSLIGRVSVALLGCTVFFSLKLGISGRLOSMPISVSARPHESD	631
Qy	635	KYTLKSSSFQKDKMVSMSHESVALITGVSRA-----	665
Db	632	SFLMKTESGNFKNDISVNRNGI--VGNIKYVLIIMLKMGCEHDPALYKSGQSATSLS	669
Qy	666	-DDSEALPR-MARATAEINVSFKQIKSLAFGPHRIEMLPEVLDGGM-----LK	713
Db	690	HASSELHKRPMOTBEARELVGMQEVNKAALGPHOYVLSLEVDSMLVQVESIFLCJM	749
Qy	714	IWTDRALTAQGLVY--DYTLIKLSVDSVTVSAD--GTRALVEATLEESACISDLVHP	769
Db	750	QW-QTLAQAEAKSCWRFVLIHLEVLQAHIFEBDINGEAEALIEALIEBAELVDESQPK	808
Qy	770	NNATDVRVTTTREVYFVMSKSG-WKITEGSV	798
Db	809	-NAKYSTIKIRIYLKKQBDGLMKFCQSDI	837

RESULT 13		
Q8DKU7	PRELIMINARY;	PRT; 673 AA.
ID Q8DKU7		
AC Q8DKU7;		
DT 01-MAR-2003 (T-EMBLrel. 23, Created)		
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)		
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)		
DE T10758 protein.		
GN <i>OryechodocusNames=cl0758;</i>		
OS <i>Oryechodocus elongatus</i> (Thermosynechococcus elongatus)		
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.		

OX NCBI_TaxID=32046;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BP-1;
 RX MEDLINE=2225144; PubMed=12240834;
 RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
 RA Wakamabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Nakazaki N.,
 RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the thermophilic cyanobacterium
 Thermosynechococcus elongatus BP-1.";
 RL DNA Res. 9:123-130(2002).
 DR EMBL; AP005371; BAC08309.1;
 DR InterPro; IPR001623; DnaJ_N.
 DR InterPro; IPR008943; TPR-like.
 DR Pfam; PF00226; DnaJ_1.
 DR SMART; SM00271; DnaJ_1.
 DR PROSITE; PS0076; DnaJ_2; 1.
 KM Complete proteome.
 SQ SEQUENCE 673 AA; 75741 MW; 3427C6E46BCB83A CRC64;

Query Match 10.4%; Score 423; DB 2; Length 673;
 Best Local Similarity 25.4%; Pred. No. 2.2e-17;
 Matches 191; Conservative 105; Mismatches 307; Indels 150; Gaps 28;

OY 86 IPIDFYVLAQTHFLTDGIRRAFEARVSKRPGFSDDLISRQILQAACETLSNPRS 145
 DB 3 IPIDFYVLAQTHFLTDGIRRAFEARVSKRPGFSDDLISRQILQAACETLSNPRS 62
 OY 146 RREYNE-----GLLDEEATVITDVPMDKPGALCVLOEGGETEIVLNGEALL 194
 DB 63 RDADVDRHCRITVDDPLAQDLPDATTPIHEISDEQSGALLILVELGNVAVNLGDAFL 122
 OY 195 KERLPKSFQ-----DVLVMAFLDVSRDAMALDPDFITGIEFEVALKLL 243
 DB 123 KKVY---FEENRPTSPAAVADITLVALATLIELGREWM--QRQSYSAASQLRAGLQVL 177
 OY 244 QEEGASLAPDLRAQIDETLEIIPRYVLELGLPLGDDYAAKGLNGSVRNILMVG 303
 DB 178 QR---VNLPELQGFQTELRNLRPIELALPLSD--SANRQIGILRKQLSERGG 232
 OY 304 -GGASALVGGITRE--KFMNEAFLEMTAEQVDFVATPSNIPAESFEVEVALVAQ 359
 DB 233 IEGGDDRSGLTYVEDPKFLIQLSHSLTVARQGLF-----BRESRPSAVATYLAH 285
 OY 360 AFIGKKEHLID----ADKQFOQLQAKWAMEIPMLYDTRNNMEIDFGLERGLCALL 414
 DB 286 ALVARGVHEIQPSYICRAKDLIQOL-----LPHQ-----DVLIELASCLLL 326
 OY 415 IGVKDEGRMVLGLDSESOYRNPAIVEFVLNSNRDNDLPGCLKLETTVLAVGPPRF 474
 DB 327 LGGQTE---ALMALHSDQPTL-DPIRHHAG-EGADRLPGLYYYTTOGLTEIIPAF 379
 OY 475 RDTFKKKFKLDYDDPMVLSTYLERVEVVOGSPALAAATMARIGAEHVKASAMQALQVF 534
 DB 380 RDLGEFTVALAEAYADANVOTYLEALSSEDSTAPRPPTTA----- 419
 OY 535 PSRYTDRNSAEPKQVQETVPSVDVGNVNGDGEQGVFIABAVPSNFETNDYAIRAGV 594
 DB 420 -----SALPEYIRPTV-AVPP-----PLSFTAEVL-PLDQSRLLGGGLSASA 459
 OY 595 SESSVDETTYEMSVADMLKEASVK-----ILAAVVA-IGLISLSQKFLKSS 642
 DB 460 FTSPATATGTSMQPSFRKRSPRNRCQKQVTFMAGAGVAVLGGAL--AKYWPAAKT 517
 OY 643 SFORKDVAWSSMESDVATIGSVRADDSBAL--PRMDARTAEINYSKKOKISLAFGPDHR 699
 DB 518 AEAAPPVYTAPIPTVATPTPOPTTIAITLTPEM----ARDRLHTQQIKAGLGPPE 573
 OY 700 IEMLPEVLDGRMLKIWTDRAAETLAQLGLVYDYTLKLKSVDSVTVS--ADGTRALVE----A 754
 DB 574 VDKLTITLAEPELSRMRSAQGLKSEGSYVWYTLKLNLEKVEKRLQDRSDRAVEVLAEVNE 633

OY 755 TLIESACISDLVHPENNATDVRYTTRVEVMS 787
 DB 634 RPYEQGTL-----RTDI-SYSDPYRYVYIT 656

RESULT 14

ID Q7VOH1 PRELIMINARY; PRT; 702 AA.

AC Q7VOH1;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Ordered locus names=PMW1287;
 OS Prochlorococcus marinus subsp. pastoris (Strain CCMP 1378 / MED4).
 OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
 OC Prochlorococcus.
 OX NCBI_TaxID=59919;
 RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22825698; PubMed=12917642; DOI=10.1038/nature01947;
 RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chait P.,
 RA Ahlgren N.A., Arellano A., Coleman M., Hauber L., Hees W.R.,
 RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
 RA Webb E.A., Zinser E.R., Chisholm S.W.;
 RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
 niche differentiation."
 RL Nature 424:1042-1047(2003).
 DR EMBL; BX572093; CAE19746.1;
 DR InterPro; IPR001623; DnaJ_N.
 DR Pfam; PF00226; DnaJ_1.
 DR SMART; SM00271; DnaJ_1.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 702 AA; 79938 MW; 325A8F3E74E558F CRC64;

Query Match 8.5%; Score 344.5; DB 2; Length 702;
 Best Local Similarity 20.5%; Pred. No. 1.5e-12;
 Matches 161; Conservative 155; Mismatches 305; Indels 163; Gaps 31;

OY 86 IPIDFYVLAQTHFLTDGIRRAFEARVSKRPGFSDDLISRQILQAACETLSNPRS 145
 DB 3 IPIDHFRILGVSPSATSEETLRFAQLRDKTPBGFFYEVLTQSRSELRLTADLTDPDS 62
 OY 146 RREYNESLDEEATVITDVPMDKPGALCVLOEGGETEIVLNGEALLK---ERLPKS 201
 DB 63 RRYVENLLINGASG---LDLSNREVAGLILMWESGSKSAFKITRKLQPPOTPALGSS 119
 OY 202 FKQDVLVMAFLDVSRDA--MALDPPDFTTGIEFEFEALKLLQEEGASLAPDLRAQI 259
 DB 120 READLTLAAL---TSRDLAIOBODRSYSNAADPLQESIQLOLRMGKLG--ELKRTL 172
 OY 260 DETLEIIPRYVLELGLPLGDDYAAKRLNGLSGVNILMVGSGASALVGGITRE--K 317
 DB 173 EBDLVSLPPIRIIDLSRLNDVDSHK--GLMTEHLIIKRG-----LEGKKSSYND 225
 OY 318 PMN---EAFLR-----MTAAEQVDFVATPSNIPAESFEVEVALVAQAFIGKKPHL 368
 DB 226 FLNQEESEFPQOIKPFLTVQDOIIDLFLIQKRGSSBA--GFLAFLSLTIAIGFARRPAK 283
 OY 369 LQADKQFOQLQAKWAME---IPMLYDTRNNMEIDFGLERGLCALLIGVDEGRMW 424
 DB 284 LFEARKILKDLNLSGLDSMPILGCLDLILAD-----VEQSSARFLSSDEKLRDW 333
 OY 425 IGLDSEDSQYRNPAIVEFVLNSNRDNDLPGCLKLETTVLAVGPPRDTKDKKFKL 484
 DB 334 L-----NNYRGEKLEAICIFCKRWLENDVYGRDIDDKIEIDL 371
 OY 485 GDYDDPMVLSTYLERVEVVOGSPALAAATMARIGAEHVKASAMQALQKVPSTRYDRNSA 544
 DB 372 DSWFEDEIREIOEFIOIE-----KKSNTYVFKSGPON-----KPIFOAESLSDSS 416
 OY 545 EPKDVQETVPSVD--PVGNVGRDG-----EPGVFIABAVPSNFETNDYAIRAGVSS 598

```

Db      417 TGRDLMSDNBEBRLPLPGVREBGGQEVIEENITYTDIIK-NKSIERYKAIER-IAE-- 472
Qy      599 VDETTYEMSVADMJKE-----ASVKIILAAGVAILGI-----SLSQKTFK 639
Db      473 -----LKFVGEALLENYRIFNKSSYLTYLAFILFAPFGLGVGVRRNNLKKPVQEKIID 527
Qy      640 SSSSPQRKDWV--SSMESD-----VATIGSVRADDSR-----ALPRMDARTAE 681
Db      528 NLSIINENKNVFPYEGANQDKKYLDNLSKIIISDNMEKVFPSGEIHTASPSLE--KIEN 585
Qy      682 IVSKWQKIKSLAFGPHRIEMLEPEVLDSGMLKIWTDRRAETAQLGLVYDYTLKLSDSV 741
Db      586 LINTWLVNKSFKLAGKEINLSKIVQDDLIIDRLKEERELD-IQKGI---YKNINANIENI 641
Qy      742 TVSADGTRALVENTLESACLSLVHPE-----NNATDVRTYTTTYEVFWMSKGV 791
Db      642 -----VLLTQTASRISVSVLKYSEKILKIDGELINETTPTPLKVKYIILGFSNWS 693
Qy      792 KITE 795
Db      694 KLVD 697

```

```

RESULT 15
Q7US57 PRELIMINARY; PRT; 653 AA.
AC Q7US57;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Hypoetical protein.
OS OrderedclucNames=SYNM1619;
OS Synecococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2825697; Pubmed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarten J.,
RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synecococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL: BX569693; CAE08134.1; -.
DR InterPro: IPR001623; DnaJ_N.
KW Complete proteome; Hypoetical protein.
SQ SEQUENCE 653 AA; 70997 MW; 4DF7926F67BEC537 CRC64;

```

Query Match 8.44; Score 340; DB 2; Length 653;

Best Local Similarity 21.54; Pred. No. 2,6e-12;

Matches 167; Conservative 120; Mismatches 317; Indels 156; Gaps 24;

```

Qy      86 IPIDYQVIGAQHTLTGDIRRAFEARVSKPQFGSDALISRQIILQACETLSNPRS 145
Db      3 LPIDHFLGLGVSPSADPASILRLQTRSDSPDDGFTHEGLQKALIHRSADILLTPSE 62
Qy      146 RREYVNEGILD-----DEEATVITDVPMWKYFGALCVLQEGGETEIVARVGBALLKERLP- 199
Db      63 RADYBAALLSLSATHPNETVGLDLAASSEVAGLILWEAGALFAQLARQGLQPPQAPA 122
Qy      200 --KSPKQDVVLVMAALAFVDVSRDAMALDPDFITGVFEFVEBALKLQEGASISLAPDLRA 257
Db      123 LSGGREADLTLIALAARDAARDEQ--QQRRYESAQQLARDGIELQRMGK---LPDQA 177
Qy      258 QIDETLEETPRVYVLELLGLPLGDDYAKKLNGISGVNIIIMSVG-----GGASALV 310
Db      178 RLQGEILDILPYRVLDLSRDLSD--ADARQOGISILDQLVDRDGLDPEGLDSETPAAM 235
Qy      311 GGLTRKEMNEAFLKMTAAEOVDLFAV--TPSNIPAESFEYEVVALVAQAFIGKKPHL 366
Db      236 GQADPESFQOIRRFITVQEOVDLFRGMFAEGSTEACLAVF---ALAAGYRRRKEPF 291

```

```

Qy      369 LQDADKQFOQLQCAKVMIEIPAMLVOTRNNWEIDFGLERGLCALLIGKVDCECRMILGLD 428
Db      292 LEQAREQLRLVASDLDPMPL-----LGCIDLLIGNVAAEASLHF--- 330
Qy      429 SEDSQYRNPAIVEFVLENSNRDNDLPLGLCKLLETLTAGVPPRPRTDKKFKGDDY 488
Db      331 ---SAIRDEELSLWLEH---PGDHLLAQCEYCRVLEEDVLPGYRDVDADAGVDLDAMF 383
Qy      489 DDEPMVSYLERVEVVOGSPPLAAATMARIGAEHVKASAMQALQKVPFSRYTDRNSAPEKD 548
Db      384 ADNDVQATYDRID-RQSAKLGSAATYTGAG-----LSSAPSDASS 423
Qy      549 VOETVPSVDPVGNVGRDGEVGFIAEAVRPSNPFETNDYAIAGVSESSVDETTYEMSV 608
Db      424 PHEALDDHLL-----PASEAPSSD-----P 444
Qy      609 ADMLKEASVKIILAAGVAILGISFSQKVPKSSSF-----QRKDWVSMES--DVA 658
Db      445 ANQRLSNRLRWLAAVLVGLVLAALAAVMLRPRETAPVVLQPEPDQDAVEPKPSAQDSA 504
Qy      659 TI---GSVRADDSBALPRM-----DARTAEIVSKWQIKSLAFGPHRIEMLEPEVL 707
Db      505 TLKPAQAILQPEBASEVAAGLQPLSDAPDDAQLEMLVQGMIDSKAQL--QGQPSQLPVVA 562
Qy      708 DGRMLKIWTDRRAETAQLGLVYDYTLKLSDSVTVSADGTR-----ALVEATLEES 759
Db      563 RQRLID-QVDR--ERSKAVAVAGTTVVKASVTSILDVASSQPRRIILQAGVAYSSTTDRS 619
Qy      760 ACSGLVHPENNATDVRTYTTTYEVFWMSKGVKIT 794
Db      620 GTVVDRTVP-----GSLTITYILGRDGDQWKLT 647

```

Search completed: June 10, 2005, 01:40:34

Job time : 104 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2005, 01:37:44 ; Search time 475 Seconds
(without alignments)
646.423 Million cell updates/sec

Title: US-10-600-070-2

Perfect score: 4063

Sequence: 1 MEALSHVIGIGLSPFQCRPLP.....YEVFWMSKMGKITEGVTLAS 801

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubppaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubppaa/US10E_PUBCOMB.pep:*

18: /cgn2_6/ptodata/1/pubppaa/US10F_PUBCOMB.pep:*

19: /cgn2_6/ptodata/1/pubppaa/US11A_PUBCOMB.pep:*

20: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*

21: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*

22: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4063	100.0	801	US-10-600-070-2	Sequence 2, Appl
2	4051	99.7	801	US-10-600-070-127	Sequence 127, Appl
3	4051	99.7	801	US-10-600-070-129	Sequence 129, Appl
4	4051	99.7	801	US-10-600-070-131	Sequence 131, Appl
5	4051	99.7	801	US-10-739-930-5771	Sequence 5771, Ap
6	2926	92.0	578	US-10-600-070-123	Sequence 123, App
7	1775.5	43.7	760	US-10-600-070-125	Sequence 125, App
8	1646	40.5	324	US-10-600-070-206	Sequence 126, App
9	1332	32.8	525	US-10-600-070-122	Sequence 122, App
10	1134	27.9	344	US-10-424-599-271849	Sequence 271849,
11	973	23.9	364	US-10-425-115-266516	Sequence 266516,

12	972	23.9	480	US-10-437-963-172416	Sequence 172416,
13	856	21.1	416	US-10-425-115-242115	Sequence 242115,
14	524.5	12.9	768	US-10-600-070-167	Sequence 167, App
15	499	12.3	204	US-10-424-599-177901	Sequence 177901,
16	498.5	12.3	789	US-10-600-070-194	Sequence 194, App
17	497	12.2	798	US-10-600-070-164	Sequence 164, App
18	497	12.2	798	US-10-600-070-165	Sequence 165, App
19	482	11.9	709	US-10-437-963-172415	Sequence 172415,
20	474.5	11.7	1157	US-10-437-963-155799	Sequence 155799,
21	442	10.9	631	US-10-600-070-5	Sequence 5, Appl
22	442	10.9	631	US-10-600-070-162	Sequence 162, App
23	435.5	10.7	191	US-10-767-701-35633	Sequence 35633, A
24	434.5	10.7	624	US-10-600-070-116	Sequence 116, App
25	431	10.6	819	US-10-600-070-171	Sequence 171, App
26	431	10.6	819	US-10-600-070-173	Sequence 173, App
27	429	10.6	714	US-10-600-070-169	Sequence 169, App
28	429	10.6	714	US-10-600-070-170	Sequence 170, App
29	425.5	10.5	652	US-10-600-070-115	Sequence 115, App
30	423	10.4	673	US-10-600-070-192	Sequence 192, App
31	413	10.2	836	US-10-600-070-190	Sequence 190, App
32	409.5	10.1	205	US-10-767-701-41038	Sequence 41038, A
33	401	9.9	716	US-10-600-070-160	Sequence 160, App
34	366	9.0	99	US-10-600-070-124	Sequence 124, App
35	359	8.8	566	US-10-600-070-119	Sequence 119, App
36	344.5	8.5	702	US-10-600-070-156	Sequence 156, App
37	339	8.3	66	US-10-600-070-100	Sequence 100, App
38	328.5	8.1	573	US-10-600-070-120	Sequence 120, App
39	327	8.0	357	US-10-600-070-117	Sequence 117, App
40	320	7.9	661	US-10-600-070-158	Sequence 158, App
41	317.5	7.8	515	US-10-600-070-121	Sequence 121, App
42	308.5	7.6	524	US-10-600-070-118	Sequence 118, App
43	288	7.1	164	US-10-437-963-110670	Sequence 110670,
44	263	6.5	66	US-10-600-070-99	Sequence 99, Appl
45	257	6.3	66	US-10-600-070-98	Sequence 98, Appl

ALIGNMENTS

RESULT 1

US-10-600-070-2

Sequence 2, Application US/10600070

Publication No. US20040139500A1

GENERAL INFORMATION:

APPLICANT: Oseeryoung, Katherine W.

APPLICANT: Vitka, Stanislaw

APPLICANT: Koksharova, Olga A.

APPLICANT: Gao, Hongbo

TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of

TITLE OF INVENTION: Use

FILE REFERENCE: NSU-08153

CURRENT APPLICATION NUMBER: US/10/600.070

CURRENT FILING DATE: 2003-06-20

NUMBER OF SEQ ID NOS: 206

SOFTWARE: PatentIn version 3.2

SEQ ID NO 2

LENGTH: 801

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-10-600-070-2

Query Match 100.0%; Score 4063; DB 16; Length 801;

Beet Local Similarity 100.0%; Pred. No. 0;

Matches 801; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEALSHVIGIGLSPFQCRPLPATTKLRSHNTSTTICSAKMDRLSDNFSSSS 60

DB 1 MEALSHVIGIGLSPFQCRPLPATTKLRSHNTSTTICSAKMDRLSDNFSSSS 60

QY 61 PAATTTATVSLPPSIDRBRHVPIDRYOVIQAOTHTLTDGIRAFKRVSKPQFG 120

DB 61 FATATTATVSLPPSIDRBRHVPIDRYOVIQAOTHTLTDGIRAFKRVSKPQFG 120

Query Match 99.7%; Score 4051; DB 16; Length 801;
Best Local Similarity 99.8%; Pred. No. 0;
US-10-600-070-127

[illegible]

NUMBER OF SEQ ID NOS: 206
SOFTWARE: Patentin version 3.2
SEQ ID NO 129
LENGTH: 801
TYPE: PR
ORGANISM: Arabidopsis thaliana
US-10-600-070-129

Query Match 99.7%; Score 4051; DB 16; Length 801;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MEALSHVIGIGLSPFQCRPLPATTCKLRSHNTSTTICSASKMADRLSDNFSTDSSSSS 60
1 MEALSHVIGIGLSPFQCRPLPATTCKLRSHNTSTTICSASKMADRLSDNFSTDSSSSS 60
61 FATATTATLVSPPSIDRPERHVPIDFYQVLAQTHFLTDGIRAFARVSKPQFG 120
61 FATATTATLVSPPSIDRPERHVPIDFYQVLAQTHFLTDGIRAFARVSKPQFG 120
121 FSDDALISRQIILOACETLSNPRSREYNEGLDDEATVITDVPMDKVPALCVLQEG 180
121 FSDDALISRQIILOACETLSNPRSREYNEGLDDEATVITDVPMDKVPALCVLQEG 180
121 FSDDALISRQIILOACETLSNPRSREYNEGLDDEATVITDVPMDKVPALCVLQEG 180
181 GETEIVLRVEBALKEKLPKSPKQDVVLAALFVDSRDAMALDPDFITGYEFVEAL 240
181 GETEIVLRVEBALKEKLPKSPKQDVVLAALFVDSRDAMALDPDFITGYEFVEAL 240
181 GETEIVLRVEBALKEKLPKSPKQDVVLAALFVDSRDAMALDPDFITGYEFVEAL 240
241 KILQEBGASSIAPDLRAQIDETLEETPRVYLBELGLPLGDDYAKKLNGLSGVNITLMS 300
241 KILQEBGASSIAPDLRAQIDETLEETPRVYLBELGLPLGDDYAKKLNGLSGVNITLMS 300
241 KILQEBGASSIAPDLRAQIDETLEETPRVYLBELGLPLGDDYAKKLNGLSGVNITLMS 300
301 VGGGASALVGLTREKFMNEAFIRMTAABQVDLFVATPSNIPAESEFEVALALVAQA 360
301 VGGGASALVGLTREKFMNEAFIRMTAABQVDLFVATPSNIPAESEFEVALALVAQA 360
361 FIGKPHLLODADKQFQLOQAKWAMEIPAMLVDTNNMEIDFGLRGICALLIGKVD 420
361 FIGKPHLLODADKQFQLOQAKWAMEIPAMLVDTNNMEIDFGLRGICALLIGKVD 420
421 CRMWLGDSDSQYRNPAIYEFVLNENRDNDLPGCLKLEFVLAQVFPFRDTRDK 480
421 CRMWLGDSDSQYRNPAIYEFVLNENRDNDLPGCLKLEFVLAQVFPFRDTRDK 480
481 KFKLGDYDDPMVLSYERVEVVGSPPLAAATMARI GAHVKASAMQLOKVPSTRYTD 540
481 KFKLGDYDDPMVLSYERVEVVGSPPLAAATMARI GAHVKASAMQLOKVPSTRYTD 540
481 KFKLGDYDDPMVLSYERVEVVGSPPLAAATMARI GAHVKASAMQLOKVPSTRYTD 540
541 RNSAEPDQVETVSVDPVGNNGRDEPGVFLAEAVRPSNFTNDYAIRAGVSESSVD 600
541 RNSAEPDQVETVSVDPVGNNGRDEPGVFLAEAVRPSNFTNDYAIRAGVSESSVD 600
541 RNSAEPDQVETVSVDPVGNNGRDEPGVFLAEAVRPSNFTNDYAIRAGVSESSVD 600
601 ETTVEMSVADMKEASVKIILAAVAIGLISFSQKYLKSSSFQORDMSMSDVATI 660
601 ETTVEMSVADMKEASVKIILAAVAIGLISFSQKYLKSSSFQORDMSMSDVATI 660
661 GSVRADSEALPRMDARTAEIVSKWOKIKSLAFGPHRIEMLEPEVLDGRMLKIWTDRAA 720
661 GSVRADSEALPRMDARTAEIVSKWOKIKSLAFGPHRIEMLEPEVLDGRMLKIWTDRAA 720
721 ETTAQLGLVYDTTLKLKSVDSVTVASADGTRALVEATLESACLSDLVHPENNATVTRTYTT 780
721 ETTAQLGLVYDTTLKLKSVDSVTVASADGTRALVEATLESACLSDLVHPENNATVTRTYTT 780
781 RYEVFMKSGWKITGESVLAS 801
781 RYEVFMKSGWKITGESVLAS 801

RESULT 4
US-10-600-070-131
Sequence 131, Application US/10600070
Publication No. US20040139500A1

GENERAL INFORMATION:
APPLICANT: OsteYoung, Katherine W.
APPLICANT: Vilha, Stanislav
APPLICANT: Koksharova, Olga A.
APPLICANT: Gao, Hong
TITLE OF INVENTION: Use
FILE REFERENCE: NSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
NUMBER OF SEQ ID NOS: 206
SOFTWARE: Patentin version 3.2
SEQ ID NO 131
LENGTH: 801
TYPE: PR
ORGANISM: Arabidopsis thaliana
US-10-600-070-131

Query Match 99.7%; Score 4051; DB 16; Length 801;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MEALSHVIGIGLSPFQCRPLPATTCKLRSHNTSTTICSASKMADRLSDNFSTDSSSSS 60
1 MEALSHVIGIGLSPFQCRPLPATTCKLRSHNTSTTICSASKMADRLSDNFSTDSSSSS 60
61 FATATTATLVSPPSIDRPERHVPIDFYQVLAQTHFLTDGIRAFARVSKPQFG 120
61 FATATTATLVSPPSIDRPERHVPIDFYQVLAQTHFLTDGIRAFARVSKPQFG 120
121 FSDDALISRQIILOACETLSNPRSREYNEGLDDEATVITDVPMDKVPALCVLQEG 180
121 FSDDALISRQIILOACETLSNPRSREYNEGLDDEATVITDVPMDKVPALCVLQEG 180
121 FSDDALISRQIILOACETLSNPRSREYNEGLDDEATVITDVPMDKVPALCVLQEG 180
181 GETEIVLRVEBALKEKLPKSPKQDVVLAALFVDSRDAMALDPDFITGYEFVEAL 240
181 GETEIVLRVEBALKEKLPKSPKQDVVLAALFVDSRDAMALDPDFITGYEFVEAL 240
181 GETEIVLRVEBALKEKLPKSPKQDVVLAALFVDSRDAMALDPDFITGYEFVEAL 240
241 KILQEBGASSIAPDLRAQIDETLEETPRVYLBELGLPLGDDYAKKLNGLSGVNITLMS 300
241 KILQEBGASSIAPDLRAQIDETLEETPRVYLBELGLPLGDDYAKKLNGLSGVNITLMS 300
241 KILQEBGASSIAPDLRAQIDETLEETPRVYLBELGLPLGDDYAKKLNGLSGVNITLMS 300
301 VGGGASALVGLTREKFMNEAFIRMTAABQVDLFVATPSNIPAESEFEVALALVAQA 360
301 VGGGASALVGLTREKFMNEAFIRMTAABQVDLFVATPSNIPAESEFEVALALVAQA 360
361 FIGKPHLLODADKQFQLOQAKWAMEIPAMLVDTNNMEIDFGLRGICALLIGKVD 420
361 FIGKPHLLODADKQFQLOQAKWAMEIPAMLVDTNNMEIDFGLRGICALLIGKVD 420
421 CRMWLGDSDSQYRNPAIYEFVLNENRDNDLPGCLKLEFVLAQVFPFRDTRDK 480
421 CRMWLGDSDSQYRNPAIYEFVLNENRDNDLPGCLKLEFVLAQVFPFRDTRDK 480
481 KFKLGDYDDPMVLSYERVEVVGSPPLAAATMARI GAHVKASAMQLOKVPSTRYTD 540
481 KFKLGDYDDPMVLSYERVEVVGSPPLAAATMARI GAHVKASAMQLOKVPSTRYTD 540
481 KFKLGDYDDPMVLSYERVEVVGSPPLAAATMARI GAHVKASAMQLOKVPSTRYTD 540
541 RNSAEPDQVETVSVDPVGNNGRDEPGVFLAEAVRPSNFTNDYAIRAGVSESSVD 600
541 RNSAEPDQVETVSVDPVGNNGRDEPGVFLAEAVRPSNFTNDYAIRAGVSESSVD 600
541 RNSAEPDQVETVSVDPVGNNGRDEPGVFLAEAVRPSNFTNDYAIRAGVSESSVD 600
601 ETTVEMSVADMKEASVKIILAAVAIGLISFSQKYLKSSSFQORDMSMSDVATI 660
601 ETTVEMSVADMKEASVKIILAAVAIGLISFSQKYLKSSSFQORDMSMSDVATI 660
661 GSVRADSEALPRMDARTAEIVSKWOKIKSLAFGPHRIEMLEPEVLDGRMLKIWTDRAA 720
661 GSVRADSEALPRMDARTAEIVSKWOKIKSLAFGPHRIEMLEPEVLDGRMLKIWTDRAA 720
721 ETTAQLGLVYDTTLKLKSVDSVTVASADGTRALVEATLESACLSDLVHPENNATVTRTYTT 780
721 ETTAQLGLVYDTTLKLKSVDSVTVASADGTRALVEATLESACLSDLVHPENNATVTRTYTT 780

QY 781 RYEFWMSKGMKITEGSLAS 801
DB 781 RYEFWMSKGMKITEGSLAS 801

RESULT 5

US-10-739-930-5771
; Sequence 5771, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739, 930
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5771
; LENGTH: 801
; TYPE: PR1
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C13643_1.P
US-10-739-930-5771

Query Match 99.7%; Score 4051; DB 16; Length 801;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MEALSHVIGLSPPQICRLPATTKLRSHNTSTTICSAKMDRLIDFNFTSDSSSS 60
DB 1 MEALSHVIGLSPPQICRLPATTKLRSHNTSTTICSAKMDRLIDFNFTSDSSSS 60
QY 61 FATTATTAATVSLPSTIDRRHVPIPIDYQVLAQTHPLTGIRAFARVSKPQFG 120
DB 61 FATTATTAATVSLPSTIDRRHVPIPIDYQVLAQTHPLTGIRAFARVSKPQFG 120
QY 121 FSDALISRRQIIQAACETSNPRSRREYEGLLDDEATYITDVPMDKPGALCVLOEG 180
DB 121 FSDALISRRQIIQAACETSNPRSRREYEGLLDDEATYITDVPMDKPGALCVLOEG 180
QY 121 FSDALISRRQIIQAACETSNPRSRREYEGLLDDEATYITDVPMDKPGALCVLOEG 180
DB 121 FSDALISRRQIIQAACETSNPRSRREYEGLLDDEATYITDVPMDKPGALCVLOEG 180
QY 181 GETEIVLRVGEALLKERLPKSFQDQVLAFLVDSRDAMALDPDFITGYEFVEAL 240
DB 181 GETEIVLRVGEALLKERLPKSFQDQVLAFLVDSRDAMALDPDFITGYEFVEAL 240
QY 241 KLIQEGGASLADPLRAQIDETLEETPRVYELGLPLGDDVYAAKRLNGLSGRNLTMS 300
DB 241 KLIQEGGASLADPLRAQIDETLEETPRVYELGLPLGDDVYAAKRLNGLSGRNLTMS 300
QY 301 VGGGASALVGGLTREKFNMEAFLRMTAAEQVDLFAVTPSNIPAESFEYEVALLVAQA 360
DB 301 VGGGASALVGGLTREKFNMEAFLRMTAAEQVDLFAVTPSNIPAESFEYEVALLVAQA 360
QY 361 FICKKPHLLQDADKQFOQLQAQVMAEIPAMLYDTNNMEIDFGELRGICALLIGKVD 420
DB 361 FICKKPHLLQDADKQFOQLQAQVMAEIPAMLYDTNNMEIDFGELRGICALLIGKVD 420
QY 421 CRWMLGLDSRDSQYRPAIYEFLENSNRDNDLPGCLKLETWLAGVFPFRDTRDK 480
DB 421 CRWMLGLDSRDSQYRPAIYEFLENSNRDNDLPGCLKLETWLAGVFPFRDTRDK 480
QY 481 KFKLGDYVDDPMTLSYLERVEVQGPLAAATWARI GAHVKASAMQALOKVFPSTYD 540
DB 481 KFKLGDYVDDPMTLSYLERVEVQGPLAAATWARI GAHVKASAMQALOKVFPSTYD 540
QY 541 RNSABPKDQVETVSVDPVGNVNGRDEPGVFLAEAVRPSENFETNDYAIRAGVSESSVD 600
DB 541 RNSABPKDQVETVSVDPVGNVNGRDEPGVFLAEAVRPSENFETNDYAIRAGVSESSVD 600
QY 601 ETTVEMSVADMILKEASVKILAAQVAILGLISLFSQKFLKSSSFORDMWSMESDVATI 660
DB 601 ETTVEMSVADMILKEASVKILAAQVAILGLISLFSQKFLKSSSFORDMWSMESDVATI 660

QY 661 GSVRADDSSEALPMDARTANIVSKWOKISLAFGPDHRIEMLPEVLGDMKLTWDRAA 720
DB 661 GSVRADDSSEALPMDARTANIVSKWOKISLAFGPDHRIEMLPEVLGDMKLTWDRAA 720
QY 721 ETRQGLVYDVTYLLKLSVDSVTVSADGTRALVEATLESACSLDVHPENNAIDVRYTT 780
DB 721 ETRQGLVYDVTYLLKLSVDSVTVSADGTRALVEATLESACSLDVHPENNAIDVRYTT 780
QY 781 RYEFWMSKGMKITEGSLAS 801
DB 781 RYEFWMSKGMKITEGSLAS 801

RESULT 6

US-10-600-070-123
; Sequence 123, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteeryoung, Katherine W.
; APPLICANT: Vilcha, Stanislav
; APPLICANT: Kokharova, Olga A.
; APPLICANT: Gao, Hongo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 123
; LENGTH: 578
; TYPE: PR1
; ORGANISM: Arabidopsis thaliana
US-10-600-070-123

Query Match 72.0%; Score 2926; DB 16; Length 578;
Best Local Similarity 99.8%; Pred. No. 2.8e-237;
Matches 577; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 79 RPRRHVPIPIDYQVLAQTHPLTGIRAFARVSKPQFGSDALISRRQIIQAAC 138
DB 1 RPRRHVPIPIDYQVLAQTHPLTGIRAFARVSKPQFGSDALISRRQIIQAAC 60
QY 139 TSNPSRSREYEGLLDDEATYITDVPMDKPGALCVLOEGGETIIVLRVGEALLKERL 198
DB 139 TSNPSRSREYEGLLDDEATYITDVPMDKPGALCVLOEGGETIIVLRVGEALLKERL 120
QY 61 TSNPSRSREYEGLLDDEATYITDVPMDKPGALCVLOEGGETIIVLRVGEALLKERL 120
DB 61 TSNPSRSREYEGLLDDEATYITDVPMDKPGALCVLOEGGETIIVLRVGEALLKERL 120
QY 199 PKSFQDQVLAFLVDSRDAMALDPDFITGYEFVEALKLIQEGGASLADPLRAQ 258
DB 199 PKSFQDQVLAFLVDSRDAMALDPDFITGYEFVEALKLIQEGGASLADPLRAQ 180
QY 121 PKSFQDQVLAFLVDSRDAMALDPDFITGYEFVEALKLIQEGGASLADPLRAQ 180
DB 121 PKSFQDQVLAFLVDSRDAMALDPDFITGYEFVEALKLIQEGGASLADPLRAQ 180
QY 259 IDETLEETPRVYELGLPLGDDVYAAKRLNGLSGRNLTMSVGGGASALVGGLTREK 318
DB 259 IDETLEETPRVYELGLPLGDDVYAAKRLNGLSGRNLTMSVGGGASALVGGLTREK 240
QY 319 MNEAFLRMTAAEQVDLFAVTPSNIPAESFEYEVALLVAQAFTGKPHLLQDADKQFOQ 378
DB 319 MNEAFLRMTAAEQVDLFAVTPSNIPAESFEYEVALLVAQAFTGKPHLLQDADKQFOQ 300
QY 241 MNEAFLRMTAAEQVDLFAVTPSNIPAESFEYEVALLVAQAFTGKPHLLQDADKQFOQ 300
DB 241 MNEAFLRMTAAEQVDLFAVTPSNIPAESFEYEVALLVAQAFTGKPHLLQDADKQFOQ 300
QY 379 LQAQVMAEIPAMLYDTNNMEIDFGELRGICALLIGYDRCRWMLGLDSRDSQYRNP 438
DB 379 LQAQVMAEIPAMLYDTNNMEIDFGELRGICALLIGYDRCRWMLGLDSRDSQYRNP 360
QY 301 LQAQVMAEIPAMLYDTNNMEIDFGELRGICALLIGYDRCRWMLGLDSRDSQYRNP 360
DB 301 LQAQVMAEIPAMLYDTNNMEIDFGELRGICALLIGYDRCRWMLGLDSRDSQYRNP 360
QY 439 IYEFLENSNRDNDLPGCLKLETWLAGVFPFRDTRDKKFKLGDYVDDPMTLSY 498
DB 439 IYEFLENSNRDNDLPGCLKLETWLAGVFPFRDTRDKKFKLGDYVDDPMTLSY 420
QY 361 IYEFLENSNRDNDLPGCLKLETWLAGVFPFRDTRDKKFKLGDYVDDPMTLSY 420
DB 361 IYEFLENSNRDNDLPGCLKLETWLAGVFPFRDTRDKKFKLGDYVDDPMTLSY 420
QY 499 RVEVQGPLAAATWARI GAHVKASAMQALOKVFPSTYD RNSABPKDQVETVSVDP 558
DB 499 RVEVQGPLAAATWARI GAHVKASAMQALOKVFPSTYD RNSABPKDQVETVSVDP 480
QY 421 RVEVQGPLAAATWARI GAHVKASAMQALOKVFPSTYD RNSABPKDQVETVSVDP 480
DB 421 RVEVQGPLAAATWARI GAHVKASAMQALOKVFPSTYD RNSABPKDQVETVSVDP 480
QY 559 VGNVNGRDEPGVFLAEAVRPSENFETNDYAIRAGVSESSVD ETTVEMSVADMILKEASVK 618
DB 559 VGNVNGRDEPGVFLAEAVRPSENFETNDYAIRAGVSESSVD ETTVEMSVADMILKEASVK 618

Db 481 VGNVVGADGPGVYIAEAVPSENFETNDYAIRAGVSSSVDETTEVMSVADMKEASVK 540
 QY 619 ILAAGVAILGILSFSQKCYFLKSSSSFORKQMVSMESD 656
 Db 541 ILAAGVAILGILSFSQKCYFLKSSSSFORKQMVSMESD 578

RESULT 7

US-10-600-070-125
 ; Sequence 125, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Viltha, Stanislav
 ; APPLICANT: Koksharova, Olga A.
 ; APPLICANT: Gao, Hong
 ; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; NUMBER OF SEQ ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 125
 ; LENGTH: 760
 ; TYPE: PRT
 ; ORGANISM: *Oryza sativa*
 US-10-600-070-125

Query Match 43.7%; Score 1775.5; DB 16; Length 760;
 Best Local Similarity 47.9%; Pred. No. 4,1e-140;

Matches 390; Conservative 119; Mismatches 213; Indels 93; Gaps 16;

QY 12 SPQOLCRPATTKLRSHNTSTTIC-SASKMADRLSDENF-----TSOSSSSSPATAT 65
 Db 14 APAPFSLPRPRPRPRPRPPHPSMAACRAASRMARLFPADFLPTAASPDPSPAPAA 73
 QY 66 TTAIVSLPSPIDRPRHVPPIPIIDFYQVLAQTHFLTDGIRRAFEAVSPRQPGFDDA 125
 Db 74 APASPFVPLPPDAERSLPLQVDFYVLAQEPFLDGIIRRAFEAVIAKPPQGYSTDA 133
 QY 126 LISRRQILQAACETLSNPRSREYNEGILDEBEATVITDPMVKVPGALCVLOEGSTEI 185
 Db 134 LVERKQMLQJAHNTLWQNSRTQYDRALSNREBALTMIDAMK-----ZAGALA 184
 QY 186 VLRVGEALLKRLPKSFQDVVLVMAALFLDVSRDAMALDPDPFITGYEFVEBALKILQ 245
 Db 185 VLVTGEQLLDRPKRPFKQDVVLAMALAYVDSLSDAMASPPDVIGCCVLERALKILO 244
 QY 246 EGASSLAPDLRAQIDETLEITPRVYELGLPGDDYAAKRLNGSGVRNIIIMSVGGG 305
 Db 245 DGASNLAPDLISQIDETLEITPRCVLELSTPIOTHHKKROGLOGARNIIMSVRG 304
 QY 306 ASALVGLTREKFNNEAFLRMTAABOVDLFAATPSNI PASFEYEVALLVAQAFIGK 365
 Db 305 IAVVGGGFSREARNEAFLMTSTEQNDPFSKTPNSIPPEWFEIYNVALHVAQAIIISK 364
 QY 366 PHLLQADAKQPOQLQAQVNAMEIPAML YDTRNMEIDFGLERGLCALLIGKDECEM 425
 Db 365 PGRIMADDLFEQLOKFNIGS---HYAYDN---EMDLALERAFGLLVGDVSKGMWL 416
 QY 426 GLDSESOYRNPAI VERVLENSN-RDNDNDLPGICLLETWLAGVPPRRDTCDDKFKL 484
 Db 417 GIDNESSPYADPKILEITVNTSSISEENDLLPGICLLETWLIPEVPRSDTRGMQFRL 476
 QY 485 GDYDDPMVLSYLERVVOGSPPLAAATWATIGAE-----HYKASAMQALQVFP-SR 537
 Db 477 GDYDDPEVLSYLERVVOGSPPLAAATWATIGAE-----HYKASAMQALQVFP-SR 536
 QY 538 YTRNSAEPDVOBTQVSPVGNVGRDGPVYIAEAVPSENFETNDYAIRAGVSS 597
 Db 537 QLRDSAMEN-----KDG-PCGYL-----ENFDQ----- 559

QY 598 SVDETTEMSVADMKEASVYKILAGVAILGILSFSQKCYFLKSSSSFORKQMVSMESD- 656
 Db 560 -----ENAPAHBSRBAALKITISGALFALLATYAKY-----LPKRPPLAIRBEH 605
 QY 657 --VATIGSVADDBEAL-----PRMDARTENIVSKQIKSLAFGPDHRIEMLEPVL 707
 Db 606 GSVAVANSVDSTDDPALDEDPVHIFRMDAKLAEDIVKQGISKALGPEHVSASLOEVL 665
 QY 708 DGRMLKXTDPAATAGLVYDTTLKLSVDSVTSVADGTRALVENTLESACLSDLVH 767
 Db 666 DGNMLKWTIDRAAIEHRRGWMEYTLSDVITISITSLDGRATVEATIDEAGOLTDVTE 725
 QY 768 PENNATDRTYTRYEVFWSK-SGMKITEGSVLAS 801
 Db 726 PRNDSTYDTKITTYEMAFSLGSGKITTEGAVALS 760

RESULT 8

US-10-600-070-206
 ; Sequence 206, Application US/10600070
 ; Publication No. US20040139500A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Osteoryoung, Katherine W.
 ; APPLICANT: Viltha, Stanislav
 ; APPLICANT: Koksharova, Olga A.
 ; APPLICANT: Gao, Hong
 ; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
 ; FILE REFERENCE: MSU-08153
 ; CURRENT APPLICATION NUMBER: US/10/600,070
 ; NUMBER OF SEQ ID NOS: 206
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 206
 ; LENGTH: 324
 ; TYPE: PRT
 ; ORGANISM: *Arabidopsis thaliana*
 US-10-600-070-206

Query Match 40.5%; Score 1646; DB 16; Length 324;
 Best Local Similarity 100.0%; Pred. No. 8.6e-130;

Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEALSHVIGISLPQOLCRPATTKLRSHNTSTTICSASKMADRLSDENFTSDSSSS 60
 Db 1 MEALSHVIGISLPQOLCRPATTKLRSHNTSTTICSASKMADRLSDENFTSDSSSS 60
 QY 61 FATATTATVSLPSPIDRPRHVPPIPIIDFYQVLAQTHFLTDGIRRAFEAVSKPQFG 120
 Db 61 FATATTATVSLPSPIDRPRHVPPIPIIDFYQVLAQTHFLTDGIRRAFEAVSKPQFG 120
 QY 121 FSDDALISRRQILQAACETLSNPRSREYNEGILDEBEATVITDPMVKVPGALCVLOEG 180
 Db 121 FSDDALISRRQILQAACETLSNPRSREYNEGILDEBEATVITDPMVKVPGALCVLOEG 180
 QY 181 GERFIVLRVGEALLKRLPKSFQDVVLVMAALFLDVSRDAMALDPDPFITGYEFVEBAL 240
 Db 181 GERFIVLRVGEALLKRLPKSFQDVVLVMAALFLDVSRDAMALDPDPFITGYEFVEBAL 240
 QY 241 KLOEBGASSLAPDLRAQIDETLEITPRVYELGLPGDDYAAKRLNGSGVRNIIIMSV 300
 Db 241 KLOEBGASSLAPDLRAQIDETLEITPRVYELGLPGDDYAAKRLNGSGVRNIIIMSV 300
 QY 301 VGGGASALVGLTREKFNNEAFL 324
 Db 301 VGGGASALVGLTREKFNNEAFL 324

RESULT 9

US-10-600-070-122
 ; Sequence 122, Application US/10600070
 ; Publication No. US20040139500A1

```

; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislav
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placetic Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ. ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 122
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-600-070-122

Query Match      32.8%; Score 1332; DB 16; Length 525;
Best Local Similarity 48.8%; Pred. No. 5.4e-103;
Matches 287; Conservative 88; Mismatches 137; Indels 76; Gaps 11;

QY 81 ERHVPIPIDYQVLTGQTHFLTDGIRAFARVSKPQGFSDALISRRQIIQAACETL 140
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 3 ERSLPQVDYKVLGAEHPFLGDSIRPAIRAKPQVGYSTDALGRRQMLQIADHTL 62

QY 141 SNPSRREYNEGILDEEATVITDVPWDKVPALCVLOEGGETEIVLRGEALIKERLPK 200
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 63 MNQNSTQVYRALSENREALITMDIAMD-----EAGBALAVLVTEBQLLDBRPK 113

QY 201 SFQDQVYVLMALAFDVSRDAMALDPDFTTGYEFVEBALKILQEBGASLAPDLRAQID 260
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 114 RFQDQVYVLMALAYVDLSRDAMASPPDVIGCEVEERALKILOEDASNIADPDLISQID 173

QY 261 ETLEETTPRVYVLELGLPGDDVAAKRLNGLSGVNLIWSVGSGGASALVGLTREKFN 320
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 174 ETLEETTPRCVLLSLPIDTEHHKQBGLOQARNLIWSVGKGIATVVGSGFSREAFNM 233

QY 321 EAFRLTAAEQVLDLFAVTPSNIPAESFEVEVALALVAQAFIGKKPHLLQDADKQFQQLQ 380
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 234 EAFRLRTSIEQOMFPSTKPSIPPEWEIYNVALAHVAQAIISRRPQFIMMADLPFQQLQ 293

QY 381 QAKVMAEIPAMLYDTRNNEIDFGLERGLCALLIGKVDRCRWLGIQDSQSOTRNPAY 440
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 294 KFNIGS-----HYAYDN-----EMDLALERAFCSLIVGVDSKCRMWLGIQDNSSPYRDKIL 345

QY 441 EPLLENSN-RDDNDLPGICLLETMLAGVPPRPRTKKKPKLGGYYVDPVLTSLER 499
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 346 EPLVTNSSISEENDLPGICLLETMLIFEVFPSPSRTRGMQFRLGIDYDDPEVLTSLER 405

QY 500 VEVVQSSPLAAATMARGAE-----HYKASAMQALQKVP--SRYTDRNSAEPKVOET 552
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 406 MEGGASHLAAALAAKLAGQATRACTVSNIAQNNKFFPLIEQDRAMENT-----460

QY 553 VFSVDVGNVNGDEBPGVFIAVAPRSENPEFTNDVIAIRGVBSSVDTEVEMKVAIDL 612
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 461 -----KDG-FGGYL-----ENFQD-----ENAPAHDS 481

QY 613 KEASVITLAAVAIGLISLFSQYFLKSSSPQRKQVSSMESDVATI 660
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 482 RNALKIIISAGALPALAVIGAKY-----LPRKRLSAIRSHSGSV 522

RESULT 10
US-10-424-599-271849
; Sequence 271849, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ. ID NOS: 285684
; SEQ ID NO 271849
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87499C.1.pcp
US-10-424-599-271849

Query Match      27.9%; Score 1134; DB 15; Length 344;
Best Local Similarity 66.4%; Pred. No. 1.3e-86;
Matches 223; Conservative 48; Mismatches 53; Indels 12; Gaps 3;

QY 16 LCRLPATY-----KLRSS--HNTSTICSAKXADRLSDPNTSDSSSSFPATAT 66
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 12 LCTPHPTTTHPTKPNKLLSSLSRGAASLSATSKWAERLIADQFLGDAAAS--TSTS 68

QY 67 TATVSLPESIDRPERHVPPIPIFYQVLTGQTHFLTDGIRAFARVSKPQGFSDAL 126
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 69 TLSPPSVPRDPPEKXYSIPDLVYRILGAEHPFLDGIIRAVEAKFSKPPQYAFNSDAL 128

QY 127 ISRRQIIQAACETLSNPSRREYNEGILDEEATVITDVPWDKVPALCVLOEGGETEIV 186
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 129 ISRRQIIQAACETLIADPPTSREYVQSLVDDEAIAITLQIPFDKVPALCVLOEGGETEIV 188

QY 187 LRVGEALIKERLPKSKRQVYVLMALAFDVSRDAMALDPDFTTGYEFVEBALKILQOE 246
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 189 LRTGGILRERLPKTERKQDVVLMALAFDVSRDAMALSPDPIAACEMLEBALKILQOE 248

QY 247 GASSLAPDLRAQIDETLEETTPRVYVLELGLPGDDVAAKRLNGLSGVNLIWSVGSGGA 306
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 249 GATSLAPDLQAOQIDETLEETTPRCVLELALPDDDEHRAAREGGLGVNIIIMAVGCGA 308

QY 307 SALVGLTREKFNMEAFLRNTAAEQVDLFPVATPSNI 342
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 309 ATIARGFTREDPFMNEAFLHMTAAEQVELFVATPSTI 344

RESULT 11
US-10-425-115-266516
; Sequence 266516, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ. ID NOS: 369326
; SEQ ID NO 266516
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_174665C.1.pcp
US-10-425-115-266516

Query Match      23.9%; Score 973; DB 16; Length 364;
Best Local Similarity 54.8%; Pred. No. 5.4e-73;
Matches 198; Conservative 54; Mismatches 79; Indels 30; Gaps 4;

QY 12 SPQQLRLPATTYKLRSHNT-----STTICSAKXADRLSDPNTSDSSSS 57
   |||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 14 APPAFSLP-----LKRSHRPPPPGPGSTCAASRWADRLFADRHLLPMAADPPMAASSSS 68

```


Db 204 DDLKDS-----LEKLAQGSVAGDAIHDSRNDAL-----KIISAGTLALFAVLGKCLPRN 254
Qy 641 SSSFORKDWVSMESDVATIGSVRADDSBA-----LPRMDARTANIYSKMKIKSLA 693
Db 255 KS-----LPALRGEGSVAVVDSIDGPPADDEPPEIIPMDAKLAEDIRRMQSIKSA 307
Qy 694 FGPDRHIREMLPEVLIDGEMLKIMTDRAAETAOLGLVYDTLLKLKSVDSVTVSADGTRALVE 753
Db 308 LGBEHVTALQEVLTGGMMLKWTDRABRIERHGMFMEYALSGVITDITVSVGRRAAVE 367
Qy 754 ATLEBSACSLDVLHPENNATDVRTYTTTRYEVFMSK-SGWMKITGCVLAS 801
Db 368 ATLEEVGRLTVDTPKNDADYDTKTYARVEMYTSRPAQWRITGEGAVLKS 416

RESULT 14

US-10-600-070-167
; Sequence 167, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oetereyoung, Katherine W.
; APPLICANT: Vitha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placitid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 167
; LENGTH: 768
; TYPE: PRF
; ORGANISM: Nostoc punctiforme
US-10-600-070-167

Query Match 12.9%; Score 524.5; DB 16; Length 768;
Best Local Similarity 24.9%; Pred. No. 1,1e-34;
Matches 212; Conservative 128; Mismatches 286; Indels 227; Gaps 28;

Qy 84 VPIPIPIFYQVLAQOTHELTGIRAFARVSKPPQFGSDALISRKQIIOAACTISNP 143
Db 1 VRIPLLYRILGRLPLAASEQOLROAVSDRIVOLPRREYSQALISSRKQLEBAVYVLSDP 60
Qy 144 RSRREYNEGL-----DDEAATVIT-----DVPMDKVGKALCVLQ 178
Db 61 KQSTYDQLYLAAYDPDNLAAVAQENRTSTKGSSTQSLGIEITDDELGALLILQ 120
Qy 179 EGGETEIVLAVGEAL--KERLPKSPKQ-----DVLVMAIAFLADVSRD- 220
Db 121 ELGELYELVLAQRPYLNVKNSATSSRSKNNLADBEIYESAHPVVLTVALLACLEGREQ 180
Qy 221 -----AMALDPDPITGYEFVEBALKLOEBGASSLAFLRAQIDTLEITPRVY 271
Db 181 MOOCHYENNAISLE-----TGOS-----LVREG---LFSSIOAEIOADLYKLAPRYI 225
Qy 272 LELLGLPLGDVNAKRLNGSGRNILMSVGS--GSAALVGGILTR--KEMNAEFLRMT 327
Db 226 LELLALP--OETAKNSQGLLELQNLLEDRGGIDGTNNDESGNITDPLFLFIQOLRNHLT 283
Qy 328 AAEQVLDLVATPSPNIPAESFEVYEVVALVAQAFTIGKPHILODPAKQFOOLOQAKVAM 387
Db 284 VAEQHLKLEAQSQR--SSAVATYLAVALIARQPAQRPALIRQARQVRLGRQ----- 337
Qy 388 EIRPAMLYTRNNKEIDPGLERGLCALIGVDEGRMMLGLDSEDSQYRNALIVEFLVENS 447
Db 338 -----DVHLEOSICALLIGQTEATRVLBLESO---YE---ALAFIREKS 376
Qy 448 NRDDNDLPEGLCLLETWLAGVVPFRPRDKKKFKGVDYDDPMVLSYLERVEVVGSP 507
Db 377 -QDSPDLLPEGLCYAEQMLQHEVFPFRFLDANQOAFKDYFANQOVAYILE----- 426

Qy 508 LAAATMARIAGAEHVKASAMQALQKVPSPRYTDRNSAEPK-----DVQETVFSV 556
Db 427 --ALPTDAQTTNEMAYINPOYFPQAKAKNTHPANNSTKTSASFNHARVPBDLPET----- 480
Qy 557 DPVGNVNGRDEGEGVIAEIVRSENFETNDYAIRAGVSSSSVDETTEVNSVMDLK--- 613
Db 481 -----PTK--ETSEYDNFSPMMWSSGSIKSEVPAERMSRGT 516
Qy 614 -----EASVKILAG----- 623
Db 517 NQHLNLSAASASAGHQKRRKPTPSASRERIPDNRPHSRRRRRRTFANTIEGKRLV 576
Qy 624 --VAIGLISL-----FSQYFLKSSSFORKDWVSMESDVATIGSVRADDS 669
Db 577 WRVFIISVILVFWLATTFGWLKNLFFQPSPDPQLFVQINQPPLPIDPNRRKPESE 636
Qy 670 ALPRMDARTENIVSKQKIKSLAFGPDRHIREMLPEVLIDGEMLKIMTDRAAETAOLGLVY 729
Db 637 EGBLVTNAE-AEEVIRHMLSTKAAALGPNNHINNLEQILTGSALSQWR-LIAQONKLDNRY 694
Qy 730 ---DYTLKLKSVDSVTVSADGTRALVEATLEBSACSLDVLHPENNATDVRTYTRYEVFW 786
Db 695 RKPDHSLIKIESVAKIGLPAD--RAAVERATYKEVTQLYENNQPFQSSND--KLARVLDLIR 750
Qy 787 SKSGWKITGCVL 799
Db 751 ERGKWRIGSTSVV 763

RESULT 15

US-10-424-599-177901
; Sequence 177901, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 177901
; LENGTH: 204
; TYPE: PRF
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_131660C.1.pep
US-10-424-599-177901

Query Match 12.3%; Score 499; DB 15; Length 204;
Best Local Similarity 51.2%; Pred. No. 1.9e-33;
Matches 104; Conservative 39; Mismatches 58; Indels 2; Gaps 2;

Qy 601 ETTVEWSVADMLEKASVKILAGVAIGLISLFSQKTF-LKSSSFORKDWVSMESDVAT 659
Db 2 ETRKGFITBEIRIGHAIVQIMCAGVIGVTLVGLKFLPRNNGSPIARKMTGSMVSDTIN 61
Qy 660 TGSVRADDS-EALPRMDARTENIVSKQKIKSLAFGPDRHIREMLPEVLIDGEMLKIMTD 718
Db 62 LGSIGDEBEKEQLPKMDARVAEALVWKQSVKSEAGPCHCLGRLEHVLIDGEMLKIMTD 121
Qy 719 AAEQAOLGLVYDVTLLKLSVDSVTVSADGTRALVEATLEBSACSLDVLHPENNATVRY 778
Db 122 AAEIARAGVSYDTLLDLNIDSVTISONGRRRAVETLLKSTHLNAVGHQPHDASNRIT 181
Qy 779 TTRYEVFMSKSGWKITGCVLAS 801
Db 182 TTRYEVFMSKSGWKITGCVLAS 204

Fri Jun 10 09:57:55 2005

us-10-600-070-2.rapb

Page 9

Search completed: June 10, 2005, 01:50:42
Job time : 477 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 9, 2005, 15:02:56 ; Search time 1188 Seconds
(without alignments)
12476.017 Million cell updates/sec

Title: US-10-600-070-3

Perfect score: 3667

Sequence: 1 tgcctgcataaaggagaac.....ctataacataaggctacaa 3667

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68473088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_esc1.*
2: gb_esc2.*
3: gb_hic.*
4: gb_esc3.*
5: gb_esc4.*
6: gb_esc5.*
7: gb_esc6.*
8: gb_gsa1.*
9: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	725.2	19.8	919	3	CNS09YUH
2	436.2	11.9	741	3	CNS09ZSS
3	434.6	11.9	741	3	CNS09YNM
4	402.2	11.0	561	1	A1998415
5	363.6	9.9	785	8	B2437564
6	359.4	9.8	434	5	BP620404
7	336.4	9.2	897	6	CD573714
8	315.8	8.6	534	5	B0834167
9	311.8	8.5	320	1	AV830764
10	307.2	8.4	320	8	B2765587
11	305.6	8.3	954	5	CG961431
12	295.6	8.1	631	9	B0046755
13	293.4	8.0	832	7	CO079829
14	265.4	7.2	751	9	CR486014
15	254	6.9	849	7	CO075595
16	252.6	6.9	769	7	CO071968
17	252	6.5	423	1	AV812946
18	239.6	6.5	703	7	CNS18842
19	237.2	6.5	703	7	CNS18842
20	212.6	5.8	521	7	CNS14655
21	212.6	5.8	593	7	CNS904734
22	210	5.7	660	4	B1268376
23	208.8	5.7	660	4	B1268376
24	194	5.3	2307	9	CL965374

25	193.2	5.3	510	7	CF603268
26	166.4	4.5	542	5	B0586205
27	161.8	4.4	542	5	CV241483
28	161.8	4.4	821	7	CK090561
29	161.8	4.4	922	7	CV264112
30	161	4.4	545	5	B0410206
31	161	4.4	692	9	CL723297
32	158	4.3	608	2	BE472035
33	157.8	4.3	722	7	CO117046
34	157.4	4.3	321	9	AL936312
35	156.2	4.3	947	9	CG088802
36	155.6	4.2	764	6	CD903230
37	153	4.2	723	7	CO079828
38	152.6	4.2	574	2	AM696905
39	152.4	4.2	611	6	CA173502
40	148.8	4.1	866	6	CD573715
41	148.2	4.0	683	7	CO076192
42	147.8	4.0	258	6	CD479481
43	147.8	4.0	914	8	B2722703
44	147.6	4.0	476	5	B0835742
45	147.4	4.0	652	9	CL584270

ALIGNMENTS

RESULT 1
CNS09YUH
LOCUS
DEFINITION
CNS09YUH 919 bp mRNA linear HTC 04-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLT892C08 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).

ACCESSION
BX841670.1 GI:42406830
VERSION
HIC; GSIT cDNA.
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quecchi,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

TITLE
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation

JOURNAL
REFERENCE
Unpublished
2 (bases 1 to 919)

COMMENT
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.

UNGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Pirap.
http://www.genoscope.cns.fr/externe/sequences/Banque_projet_EF/Full
length

FEATURES
location/Qualifiers

1..919
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"

QY 2970 TTCTCTCTGATTTTCTTTTCATTTAGGCTCAGTCAGACGATTCAGAAAG 3029
 DB 176 TTCTCTCATGAAATCGATCGCTACATAGGCTCAGTCAGACGATTCAGAAAG 235
 QY 3030 CACTTCCAGAAATGAGTCTGAGTCAGAGAAATATGATTCAGAGTCAGAAATTA 3089
 DB 236 CACTTCCAGAAATGAGTCTGAGTCAGAGAAATATGATTCAGAGTCAGAAATTA 295
 QY 3090 AGTCTGAGCTTTGGGCTGATACCGCATAGAAATGTTCCAGAGTCAGAAATTA 3149
 DB 296 AGTCTGAGCTTTGGGCTGATACCGCATAGAAATGTTCCAGAGTCAGAAATTA 340
 QY 3150 TCTCAATTCATCAATGTTGTGAACCTGTGACATGATTAATGTCGTGCTTGT 3209
 DB 341 ----- 340
 QY 3210 TGATTCGTATTTATTAAGTTTGGATGGCGAATGCTGAAGATTGGACTCAGAGCA 3269
 DB 341 -----AGTCTGAGTCAGAGCAATGCTGAAGATTGGACTCAGAGCA 384
 QY 3270 GCTGAACTGGGAGCTTGGGCTGATTAATGATTAATACATGTAATCTGTTGAC 3329
 DB 385 GCTGAACTGGGAGCTTGGGCTGATTAATGATTAATACATGTAATCTGTTGAC 444
 QY 3330 AGTGTGACAGTCTGACAGATGGAACCGTCTCTGTGTGAAGCACTGTGAGAGTCT 3389
 DB 445 AGTGTGACAGTCTGACAGATGGAACCGTCTCTGTGTGAAGCACTGTGAGAGTCT 504
 QY 3390 GCTTGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGCAAACTTACA 3449
 DB 505 GCTTGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGCAAACTTACA 564
 QY 3450 ACAAGATACGAAGTTTCTGCTCAAGTCAAGGTGGAACCACTGAAAGCTCTGTTCT 3509
 DB 565 ACAAGATACGAAGTTTCTGCTCAAGTCAAGGTGGAACCACTGAAAGCTCTGTTCT 624
 QY 3510 GCATCATATATATCTCATATGATGATGCTGAGCTTGCAGATTTCTTTGTTCTGTA 3569
 DB 625 GCATCATATATATCTCATATGATGATGCTGAGCTTGCAGATTTCTTTGTTCTGTA 684
 QY 3570 ATTCTCTCTTAAGTTAGTTTATTAATGAACAACAAAAATTAACGTTTGGCA 3626
 DB 685 ATTCTCTCTTAAGTTAGTTTATTAATGAACAACAAAAATTAACGTTTGGCA 741
 RESULT 3
 CDS09YNM 741 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSTSL592A05 of Siliques of strain col-0 of Arabidopsis thaliana
 (thale cress).
 ACCESSION BX833489.1 GI:42455179
 VERSION HTC; GSTL cDNA.
 KEYWORDS Arabidopsis thaliana
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana (thale cress)
 REFERENCE
 AUTHORS Menard M., Cruaud C., Quetier F., Scarpell C., Schachter V.,
 Temple G., Caboche M., Weissenbach J. and Salanoubat M.
 TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 741)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

COMMENT
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URVU INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
 FEATURES
 source
 1..741
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSTSL592A05"
 /csize_type="Siliques"
 /plasmid="pCMVSPORT_6"
 1..741
 /gene="At5g42480"
 ORIGIN
 Query Match 11.9%; Score 434.6; DB 3; Length 741;
 Best Local Similarity 83.3%; Pred. No. 8.7e-101;
 Matches 547; Conservative 0; Mismatches 19; Indels 91; Gaps 1;
 QY 2970 TTCTCTCTGATTTTCTTTTCATTTAGGCTCAGTCAGACGATTCAGAAAG 3029
 DB 176 TTCTCTCATGAAATCGATCGCTACATAGGCTCAGTCAGACGATTCAGAAAG 235
 QY 3030 CACTTCCAGAAATGAGTCTGAGTCAGAGAAATATGATTCAGAGTCAGAAATTA 3089
 DB 236 CACTTCCAGAAATGAGTCTGAGTCAGAGAAATATGATTCAGAGTCAGAAATTA 295
 QY 3090 AGTCTGAGCTTTGGGCTGATACCGCATAGAAATGTTCCAGAGTCAGAAATTA 3149
 DB 296 AGTCTGAGCTTTGGGCTGATACCGCATAGAAATGTTCCAGAGTCAGAAATTA 340
 QY 3150 TCTCAATTCATCAATGTTGTGAACCTGTGACATGATTAATGTCGTGCTTGT 3209
 DB 341 ----- 340
 QY 3210 TGATTCGTATTTATTAAGTTTGGATGGCGAATGCTGAAGATTGGACTCAGAGCA 3269
 DB 341 -----AGTCTGAGTCAGAGCAATGCTGAAGATTGGACTCAGAGCA 384
 QY 3270 GCTGAACTGGGAGCTTGGGCTGATTAATGATTAATACATGTAATCTGTTGAC 3329
 DB 385 GCTGAACTGGGAGCTTGGGCTGATTAATGATTAATACATGTAATCTGTTGAC 444
 QY 3330 AGTGTGACAGTCTGACAGATGGAACCGTCTCTGTGTGAAGCACTGTGAGAGTCT 3389
 DB 445 AGTGTGACAGTCTGACAGATGGAACCGTCTCTGTGTGAAGCACTGTGAGAGTCT 504
 QY 3390 GCTTGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGCAAACTTACA 3449
 DB 505 GCTTGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGCAAACTTACA 564
 QY 3450 ACAAGATACGAAGTTTCTGCTCAAGTCAAGGTGGAACCACTGAAAGCTCTGTTCT 3509
 DB 565 ACAAGATACGAAGTTTCTGCTCAAGTCAAGGTGGAACCACTGAAAGCTCTGTTCT 624
 QY 3510 GCATCATATATATCTCATATGATGATGCTGAGCTTGCAGATTTCTTTGTTCTGTA 3569
 DB 625 GCATCATATATATCTCATATGATGATGCTGAGCTTGCAGATTTCTTTGTTCTGTA 684
 QY 3570 ATTCTCTCTTAAGTTAGTTTATTAATGAACAACAAAAATTAACGTTTGGCA 3626
 DB 685 ATTCTCTCTTAAGTTAGTTTATTAATGAACAACAAAAATTAACGTTTGGCA 741

[illegible]

QY	3210	TGATTCGTATTATTATAGTTTTGGATGGGGGAAATGCTGAAGATTGTGACCTGACAGAGCA	3269
Db	369	-----AGTTTTGGATGGGGCAATGCTGAAGATTGTGACCTGACAGAGCA	326
QY	3270	GCTGAAACTCGCAGCACTGGGGTTGGTTTATGATTATACACTGTTGAAACTATCTGTTGAC	3329
Db	325	GCTGAAACTCGCAGCACTGGGGTTGGTTTATGATTATACACTGTTGAAACTATCTGTTGAC	266
QY	3330	AGTGTGACAGTCTTCAGCAAGATGAAACCCGTCCTCGGTGGAAGCAACTTCGAGAGAACTT	3389
Db	265	AGTGTGACAGTCTTCAGCAAGATGAAACCCGTCCTCGGTGGAAGCAACTTCGAGAGAACTT	206
QY	3390	GCTTGTATCTGATTTGGTTGATTCACGAAACAATGCTACTGATGTGAGAACTTACACA	3449
Db	205	GCTTGTATCTGATTTGGTTGATTCACGAAACAATGCTACTGATGTGAGAACTTACACA	146
QY	3450	ACAAGATACGAAGTTTCTGGTCCAACTCAGGGTGGAAAAATCACTGAAGCTCTGTTCTT	3509
Db	145	ACAAGATACGAAGTTTCTGGTCCAACTCAGGGTGGAAAAATCACTGAAGCTCTGTTCTT	86
QY	3510	GCATCATATATATCTCATATATGACATGTCTGAAGTTGCGAGATTCCTTTGTGTTCTGTA	3569
Db	85	GCATCATATATATCTCATATATGACATGTCTGAAGTTGCGAGATTCCTTTGTGTTGTA	26
QY	3570	ATTCTCTCTCAAGTAGTGTAT	3594
Db	25	ATTCTCTCTCAAGTAGTGTAT	1

RESULT	5
BZ437564/c	
LOCUS	785 bp DNA linear GSS 13-DEC-2002
DEFINITION	BONRN72Tf BO_1.6.2_KB_tct Brassica oleracea genomic clone BONRN72.
ACCESSION	BZ437564
VERSION	BZ437564.1 GI:26691135
KEYWORDS	GSS.
SOURCE	Brassica oleracea
ORGANISM	Brassica oleracea
AUTHORS	Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica. 1 (bases 1 to 785)
TITLE	Town,C.D., Van Aken,S., Utteebach,T., Koo,H. and Fraser,C.M.
JOURNAL	Whole genome shotgun sequencing of Brassica oleracea Unpublished (2001)
COMMENT	Other_GSSs: BONRN72Tf

9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@ligr.org
DNA is from a doubled haploid provided by Tom Osborn
Seq primer: TR
Class: sheared ends.

```

FEATURES
    source
        location/Qualifiers
            1..785
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /strain="TO1000DH3"
                /db_xref="taxon:3712"
                /clone="BONNR72"
                /clone_1lb="BO_1.6.2_KB_tot"
                /note="vector: pHOSt1; Site1
total DNA inserted into pHOSt1 using BstXI linkers"

```

ORIGIN	
Query Match	9.9%; Score 363.6; DB: 8; Length 785;
Best Local Similarity	73.3%; Pred. No. 2.1e-82;
Matches 582; Conservative	0; Mismatches 109; Indels 103; Gaps 5.

```

Qy 1061 CTAGTGAAGTTCCTGGGGCTCTCTGTATATGCAAGAGTGGTGAAGTGAAGTGAAGT 1120
Db 699 CAATGAAGTTCCTGGGGCTCTCTGTATATGCAAGAGTGGTGAAGTGAAGTGAAGT 640
Qy 1121 CTTCGGGTGGTGAAGTCTCTGTATATGCAAGAGTGGTGAAGTGAAGTGAAGT 1180
Db 639 CTTCGGGTGGTGAAGTCTCTGTATATGCAAGAGTGGTGAAGTGAAGTGAAGT 580
Qy 1181 GTTTAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1240
Db 579 GTTTAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 520
Qy 1241 CCTGATTTATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1300
Db 519 CCTGATTTATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 471
Qy 1301 TGACTGCTTGTGTAATTTGACGAGCGTGTATTAAGAACTTTCTTGAATTTGAT 1360
Db 470 ----- 471
Qy 1361 TTGTATTTAGTCTTGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1420
Db 470 -----TGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 427
Qy 1421 CAATTTGATGAGCTTTGGAAGAGATCACTCCGC-GTTATGCTTGAAGTGAAGTGAAGT 1479
Db 426 CAGTTTATGAGCTTTGGAAGAGATCACTCCGCCTTATGATGATGATGATGATGATGAT 367
Qy 1480 ACCGCTTGTGATGATTAAGCTGCGAAAGAACTTAATGTTTAAAGCGTGTGCGAATAT 1539
Db 366 ACCGCTGCGGATGAT-----AAACGACAAAGATGTTTAAAGCGTGTGCGAATAT 316
Qy 1540 TTGTGTCTGTGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1599
Db 315 CTGTGTCTGTGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 256
Qy 1600 GTTATGATGAGAGAGTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1659
Db 255 ATTATGATGAGAGAGTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196
Qy 1660 CTTTATTTATTTCTTGAAGAGATTAATTAAGTGTCTCAATTTAA-----TGATG 1714
Db 195 TTTTATTTCTTGAAGAGATTAATTAAGTGTCTCAATTTAA-----TGATG 136
Qy 1715 TTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1774
Db 135 TTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 76
Qy 1775 AAGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1833
Db 75 AAGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 16
Qy 1834 CTTTACGAGATGC 1847
Db 15 CTTTACGAGATGC 2

```

RESULT 6
 BP620404/c 434 bp mRNA linear EST 26-JUN-2004
 LOCUS BP620404 RAPL16 Arabidopsis thaliana cDNA clone RAPL16-37-C18 3',
 DEFINITION mRNA sequence.
 ACCESSION BP620404
 VERSION BP620404.1 GI:49271586
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (chale crese)
 ORGANISM Arabidopsis thaliana
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 434)
 REFERENCE Seki M., Narusawa M., Kamiya A., Ishida Y., Satou M., Sakurai T.,
 Nakajima M., Enju A., Akiyama K., Oono Y., Muramatsu M.,

```

TITLE Hayashizaki Y., Kawai J., Carninci P., Itoh M., Ishii Y.,  

JOURNAL Arakawa T., Shibata K., Shinagawa A. and Shinzaki K.  

MEDLINE Functional annotation of a full-length Arabidopsis cDNA collection  

PUBMED Science 296 (5565), 141-145 (2002)  

COMMENT Contact: Motoaki Seki  

Plant Functional Genomics Research Group  

RIKEN Genomic Sciences Center  

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  

Tel: 81-298-36-4359  

Fax: 81-298-36-9060  

Email: msek@rtc.riken.go.jp  

reversed clone; please visit our web site  

(http://efgweb.gsc.riken.go.jp/) for further details.  

FEATURES  

Location/Qualifiers  

1..434  

/organism="Arabidopsis thaliana"  

/mol_type="mRNA"  

/db_xref="taxon:3702"  

/clone="RAPL16-37-C18"  

/lab_host="DH10B"  

/clone_id="RAPL16"  

/note="Site_1: BamHI; Site_2: SalI; dark-grown"  

ORIGIN  

Query Match 9.8%; Score 359.4; DB 5; Length 434;  

Beet Local Similarity 98.0%; Pred. No. 2.2e-81;  

Matches 385; Conservative 0; Mismatches 6; Indels 2; Gaps 2;  

Qy 3224 ATAGTTTGGATGGGCGAATGCTGAAGATTTGAGTGAAGAGAGAGAGAGAGAGAGAG 3283  

Db 402 AGAGTTTGGATGGGCGAATGCTGAAGATTTGAGTGAAGAGAGAGAGAGAGAGAGAG 343  

Qy 3284 GCTTGGGTGGTGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3343  

Db 342 GCTTGGGTGGTGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 283  

Qy 3344 AGCAGATGGAACCCGCTGCTGTGTGAAGCACTGTGAAGAGTCTGTGTCTATCTGA 3403  

Db 282 AGCAGATGGAACCCGCTGCTGTGTGAAGCACTGTGAAGAGTCTGTGTCTATCTGA 223  

Qy 3404 TTTGATTCATCCAGAAACCAATGCTATGATGTGAGAACTTACACAAACAAATGAGAGT 3463  

Db 222 TTTGATTCATCCAGAAACCAATGCTATGATGTGAGAACTTACACAAACAAATGAGAGT 164  

Qy 3464 TTTGTGTCCAGTCAAGGTGGAATCA-CTGAAGGCTGTCTTGTGATCATTAATA 3522  

Db 163 TTTGTGTCCAGTCAAGGTGGAATCAAGTGAAGGCTGTCTTGTGATCATTAATA 104  

Qy 3523 CTCAATGATGAGATGCTGAGCTTGGAGATTTCTTTTGTGTAATTTCTCTCTTA 3582  

Db 103 CTCAATGATGAGATGCTGAGCTTGGAGATTTCTTTTGTGTAATTTCTCTCTTA 44  

Qy 3583 GTTAGTGTATAATGAACAAATAATTA 3615  

Db 43 GTTAGTGTATAATGAACAAATAATTA 11

```

RESULT 7
 CD573714 897 bp mRNA linear EST 12-JUN-2003
 LOCUS UCRP101_01_P12 T3 Poncirus trifoliata CTV-challenged cDNA library -
 DEFINITION UCR Poncirus trifoliata cDNA clone UCRP101_01_P12, mRNA sequence.
 ACCESSION CD573714
 VERSION CD573714.1 GI:31669616
 KEYWORDS EST.
 SOURCE Poncirus trifoliata
 ORGANISM Poncirus trifoliata
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
 1 (bases 1 to 897)

VERSION	BZ76587.1	GI:28938140
KEYWORDS	GSS.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi.	
REFERENCE	Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheek,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,I., Shim,P., Zimmerman,J. and Ecker,J.R. A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome Unpublished (2001) Contact: Joseph R. Ecker Salk Institute Genomic Analysis Laboratory (SIGAL) The Salk Institute for Biological Studies 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA Tel.: 858 453 4100 x1752 Fax: 858 558 6379 Email: eckere@salk.edu This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of AT5g42480. Class: TDNA tagged.	
JOURNAL	Location/Qualifiers	
COMMENT	1..320	
FEATURES	/organism="Arabidopsis thaliana" /mol_type="Genomic DNA" /ecotype="Col-0" /db_xref="taxon:3702" /clone="SALK_131837.28.70.x" /note="lib= "Arabidopsis thaliana TDNA insertion lines" /clone_nlb= "Note:PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tDNA_protocols.html"	
ORIGIN	Query Match 8.4%; Score 307.2; DB 8; Length 320;	
	Best Local Similarity 97.5%; Pred. No. 6,4e-68;	
	Matches 312; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
DQ	343 TTTTCAGTAGCTGTGATGTTTTTGATTATACTTAACCAAAATCAAAATTCAT	402
DB	1 TTTTCAAGTAGCTGGATGATGTTTTTGATTGAATTGACTTAATCAATCAAAATTCAT	60
DQ	403 AAACCCTTAGAGCACAACAAGCTCTTCATATATGAAAACAGAACAAGTTTTGTAGTA	462
DB	61 AAAACCTTAGAGCACATACAGTCTCTTCACTATGAAAAACAGAACAAAGTTTTGTAGTA	120
DQ	463 GCCTAAAAAAGACACTCCCATCGAAGCTCTGAGTCACGTCGCGCATTTGCTCCCCCAATTC	522
DB	121 GCCTAAAAAAGACACTCCCATCGAAGCTCTGAGTCACGTCGCGCTTTGTCTCTCCCATTTC	180
DQ	523 CAAATTATGCCGATTACCAACGGCGACGCACAAAGCTCCGACGTAGGCACAAACCTCTTACA	582
DB	181 CGATTATGCCGATTATCAACCGCGCGACGCACAAAGCTCCGACGTAGGCACAAACCTCTTACA	240
DQ	583 ACTATCTGCTCTCCGCGACGAAATGGGCGCAGCGTCTTCTCTCGCACTTCAATTTACCTTC	642
DB	241 ACTATCTGCTCTCCGCGACGCAAAATGGGCGCAGCGTCTTCTCTCGCACTTCAATTTACCTTC	300
DQ	643 GATTCCTCTCTCTCTCTCTCT 662	
DB	301 GATTCCTCTCTCTCTCTCTCT 320	
RESULT 11		
CDS	CG961431/c	954 bp DNA linear GSS 15-DEC-2003
LOCUS	CG961431	

DEFINITION	MBEKH287FB_mch2 Medicago truncatula genomic clone 73F8, genomic survey sequence.
ACCESSION	CG961431
VERSION	CG961431.1
KEYWORDS	GI:39883077
SOURCE	GSF.
ORGANISM	Medicago truncatula (barrel medic)
REFERENCE	Medicago truncatula
AUTHORS	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoid 1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
JOURNAL	1 (bases 1 to 954)
COMMENT	Town, C.D., Sherry, J., Koo, H. and Feldblyum, T.F. Sequencing of BAC ends from Medicago truncatula Unpublished (2003)
	Other_GSSs: MBEKH287FB
	Contact: Chris Town
	TIGR
	9712 Medical Center Drive, Rockville, MD 20850, USA.
	Tel: 301-838-3523
	Fax: 301-838-0208
	Email: cdtown@tigr.org
	Seq primer: CAGGAACAGCTATGACC
	Class: BAC ends.
FEATURES	Location/Qualifiers
source	1..954
	/organism="Medicago truncatula"
	/mol_type="genomic DNA"
	/culivar="genotype A17"
	/db_xref="taxon:3880"
	/clone="73F8"
	/clone_1lb="mch2"
	/note="Vector: pRelBAC11, Site_1: HindIII, Site_2: HindIII; Cook, D.R. and Klm, D.J, unpublished"
ORIGIN	
Query Match	8.3%; Score 305.6; DB: 9; Length 954;
Best Local Similarity	65.8%; Pred. No. 2.3e-67;
Matches 563; Conservative	0; Mismatches 239; Indels 54; Gaps 6
Db	1377 TGTAGAGGAAGAGCAGTAGACCTTCACCGGATTTACGTGCAAAATGATGACACTT 1436
Db	856 TGTAGGAAGAAAGGGGCAAGAGCCTTAGACCGGATTTACAAACAAATGATGAGACAC 797
QY	1437 TGGAAAGATCATCTCCGCGTTATGTCTTGAGAGCTACTTGGCTTACCGCTTGATGATT 1486
Db	796 TTGAAGAGATPAACCTCAGCTTGATATTAGAACTTTAGCCCTTCCTTATGATGAC 737
QY	1497 ACGGTGGAAGAAAGCTAAATGATTAAAGCGGTGCGGAATATTTGATGCTGTGGAG 1556
Db	736 ATCAGAGCGGAGAGAGAAAGGTCTCCAGGTGTCCGCAATCTGTGGCAATTGGAG 677
QY	1557 GAGGTGAGCATCAGCTCTGTGTGGGGGTTTGACCCGCGAGAGTTTATGATGAGCGT 1616
Db	676 GTGAGGAGACGACGCAATGTCTGGAGTTTCAACGTGAGGATTTATGATGATGAGCAT 617
QY	1617 TTTTACGAATGACAGCTGCTGAGCAGGATPACAGTTAGATA-----CCTTTTAT 1670
Db	616 TCCTGCATATGAAAGCTGCGGACAGGTGTTCATAGATATATGTGCATGTTTATTT 557
QY	1671 TTCTTTAGCATGATATATCTTTAGTTTCTTCATTTTAATGAT----- 1723
Db	556 TTAAATTTCTTCGCCCCCTTAATTTTTTGTATTAAGAGTTTACATTTATCTCTTCA 497
QY	1714 -----GTTGTGTAGTGTGATCTTTTGTAGCTACCCAG 1751
Db	496 TTAAATCTTTTATCTCAATCTTTTATTTTATGATGAACTTTTGTAGCAACCTAG 437
QY	1752 CAATATTCAGACAGATCATTTGAAGTTTACGAAGTTGACCTTGTGCTCAAGC 1811
Db	436 CAATATTCAGCTGAAGAGTTTGAAGGCTTATGAGGTTGACACTTCATGTGTTGCAAGC 377
QY	1812 TTTTATTTGTAAGAGCACACTTTTACAGAGATGCTGATAGCAATTCAGCAACTTCA 1871

Db 376 CTTTGAGTAAAGCCATCTTATCCAGATGCTGATATTTATTCATCAACTCA 317
 QY 1872 GCAGGCTTAAGTATGCGTATGAGATTCCTGCGATGTTGATGATAC-----ACGGA 1925
 Db 316 ACAACCTAAGGTAAACAATATGAGAAAGCTCCCTGTTTATCTCCATGAGATGA 257
 QY 1926 TAATTGGAGATGACTTCGGTCTAGAAAGGGAAGCTGAGCAGTCTTATAGGCAAGT 1985
 Db 256 GAAAGAGAACTTATTTGCAATTAGAAAGGGGCTGTGTGCACTGCTTGTGGAGCT 197
 QY 1986 TGATGAATGCCGATATGTTGGCTTGAAGACGTGA--GGATTCAAAATATAGAAATCCAG 2044
 Db 196 TGATCAATGTGATCAGATGTTGGGCTGATGATGACAGTCTCACTTATGAAACCAT 137
 QY 2045 CTATTTGAGACTTTGTTTGGAGAAATTCGATGATG-ACAATGATGATCTCCCTGGA 2103
 Db 136 CTATTATAGACTTTATATATGAGAAACGCAAAAGGTATGAAATGTCAGTATCTTCTGGA 77
 QY 2104 CTATGCAATTTGTTGAAACCTGATGGCAGGGGTTGCTTTC-CTAGGTTCAAGACAC 2162
 Db 76 CTCTGTAATTTGTTGAGACATGTTGATGAGATGTTTCTCTAGGTTTATGATAC 17
 QY 2163 CAAAGTAAAAAATTT 2178
 Db 16 TAAAGAGCAAACTTT 1

RESULT 12
 BU046755 631 bp mRNA linear EST 26-AUG-2002
 LOCUS BU046755
 DEFINITION Prunus persica peach developing fruit mesocarp Prunus persica cDNA
 clone PE_LBA0027104f, mRNA sequence.
 ACCESSION BU046755
 VERSION BU046755.1 GI:22486832
 KEYWORDS EST.
 SOURCE Prunus persica (peach)
 ORGANISM Prunus persica

REFERENCE
 AUTHORS Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Rosales; Rosaceae; Amygdaloidae; Prunus.
 1 (bases 1 to 631)
 CALLAHAN, A., PALMER, M., MAIN, D., WING, R. and ABBOTT, A.
 JOURNAL Peach Model Genome for Rosaceae
 COMMENT Unpublished (2002)
 CONTACT: Abbott, A.
 Dept of Genetics and Biochemistry
 Clemson University
 122 Long Hall, Clemson University, Clemson, SC 29634, USA
 Tel: 864 656 3060
 Fax: 864 656 6879
 Email: aalbert@clemson.edu

FEATURES
 SOURCE
 High quality sequence stop: 631.
 Location/Qualifiers

1..631
 /organism="Prunus persica"
 /mol_type="mRNA"
 /cultur="Loring"
 /db_xref="taxon:3760"
 /clone="PP_LBA0027104f"
 /issue_type="Mesocarp"
 /lab_host="E. coli"
 /clone_id="Peach developing fruit mesocarp"
 /note="Vector: pBluescript II SK(-); Site 1: EcoRI;
 Site 2: XhoI; authority=Prunus persica L. Batsch: The
 sequence has been trimmed to remove vector sequence and
 contains a minimum of 100 bases of phred value 20 or
 above. For more details on library preparation and
 sequence analysis go to
 http://www.genome.clemson.edu/projects/peach. To order
 this clone go to http://www.genome.clemson.edu/orders"

ORIGIN

Query Match 8.1%; Score 295.6; DB 5; Length 631;
 Best Local Similarity 71.7%; Pred. No. 7.8e-65;
 Matches 407; Conservative 0; Mismatches 149; Indels 12; Gaps 1;

QY 1708 ATGTATGTTGTGTGATGTTGATCTTTTGTATCTACCCCAAGCAATATTCAGACAG 1767
 Db 64 ATGATCTCAGCTGAGCAGGTTGATTTATTTGTACTACCCCAAGCAATATTCAGACAG 123
 QY 1768 TCATTGAAGTTTACGAAGTTGCACTGCTCTTGTGCTCAAGCTTTTATTTGTAAG 1827
 Db 124 AGCTTTAAGTTTATGAGGGTGGCTTTCGCTGTTTCTCAAGCTTTTGTGTAAGAA 183
 QY 1828 CCACACCTTTTACAGAGATGCTGATTAACCAATTCAGAACTTCAGAGGCTAAGTATG 1887
 Db 184 CTTATCAATTCAGAGATGCTGAAACCTATTCAGAACTTCAGAGCTTAAGTATCA 243
 QY 1888 GCTATGAGATTCCTCGATGTTGTATGATACAGGAATATTTGGAGATAGACTTCGGT 1947
 Db 244 GCTGTAGACATTTCTTGAACAATATATACCAAGAAAGCAGTGAATGACTTGTCT 303
 QY 1948 CTAGAAAGGGGACTCTGTGCACTGCTTATGAGCAAAAGTTGATGAATCCGTATGTG 2007
 Db 304 TTGAGAGAGGGGACTCTGTTCACCTTCTAGGGGACCTTGATGACAGTCTTCGTG 363
 QY 2008 GGCTTACACATGAGATTCACAAATATAGAAATCCAGCTATTTGGATTTGTTGGAG 2067
 Db 364 GGCTTACACATGATATGATTCACATATGAAATCATCTGTGTGATGCTTGTCTGGAG 423
 QY 2068 AATCAATGCTGATG-----ACAATGATGATCTCCCTGAGCTATGCAAAATTG 2115
 Db 424 AATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483
 QY 2116 TTGGAACCTGTTGGAGAGGGGTTGCTTCTTATGTTGATGATGATGATGATGATGAT 2175
 Db 484 TTGGAACCTGTTGGAGAGGGGTTGATGATGATGATGATGATGATGATGATGATGATGAT 543
 QY 2176 TTTTAACTCGGGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2235
 Db 544 TTCAAGCTGGAGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 603
 QY 2236 TTAGTTCAGGGTTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2263
 Db 604 GGCATTAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 631

RESULT 13
 CO079829 832 bp mRNA linear EST 15-JUN-2004
 LOCUS CO079829
 DEFINITION GR_Ea42007.r GR_Ea Gossypium raimondii cDNA clone GR_Ea42007 3',
 mRNA sequence.
 ACCESSION CO079829
 VERSION CO079829.1 GI:48749310
 KEYWORDS EST.
 SOURCE Gossypium raimondii
 ORGANISM Gossypium raimondii

REFERENCE
 AUTHORS Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 832)
 KIM, H., YU, Y., KUDRNA, D., HATFIELD, J., STUM, D., MUELLER, C.,
 UDALL, J. A., RAPP, R. A., WENDEL, J. F., RAO, K., SODERLUND, C. and
 WING, R. A.

TITLE Global assembly of Cotton ESTs
 JOURNAL Unpublished (2004)
 COMMENT Contact: Rod A. Wing
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: http://genome.arizona.edu

Plate: 42 row: 0 column: 07.

FEATURES
Location/Qualifiers

1. .832
/organism="Geosyrium raiondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_B42007"
/issue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ba"
/note="Vector: pCMV.SPORT-6.1; Site_1: NotI; Site_2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 8.0%; Score 293.4; DB 7; Length 832;

Best Local Similarity 69.2%; Pred. No. 3.2e-64;

Matches 424; Conservative 0; Mismatches 171; Indels 18; Gaps 1;

QY 1757 TTCGAGAGAGTCATTGAAGTTACGAAGTGCATTGCTCTGTGGCTCAAGCTTTTA 1816
Db 1 TTCTCGAGAAAGTTGGAAGTTTATGTGTAGACATTGCACTTGTCTCAAGCTTTT 60

QY 1817 TTGGTAAGAACCCACACCTTTTACAGAGTGTATTAAGCAATTCAGCAAGCTTCAGCAGG 1876
Db 61 TAAACAAGAAACCTCATCTCATAGAGATGCTGATTAACCTTTCAGCAGCTTCAGCAGA 120

QY 1877 CTAAAGTAATGGCTATGAGATTCCTCGATGTTGTATGATACACGGAATTAATTTGGAGA 1936
Db 121 CTAAGGTAAACACTTGTGAAGACAGTGTCTCTCTATAGCTCTGTGCGAAACCGTAGA 180

QY 1937 TAGACTCGCTTACGAAGGGAGCTGTGCACTGTTATAGCAAAAGTATGATATGCC 1996
Db 181 TAGACTTGTCTTGAAGAGGGGTCTCTGTTCATTGCTCTGTGGGAGCTTGATATGACC 240

QY 1997 GTATGTGTTGGGCTTACAGTGAAGATTCAAAATTAAGAAATCAAGTATTTGTGAGT 2056
Db 241 GTTCGTGTGGGCTTACAGTGAAGTATGCTCCCTTATTAAGAAATATCATTTGTAGAA 300

QY 2057 TTGTTTGGAGAAATCAATTCGTATGATCAATATATTCCTCGACTATGCAATTTGT 2116
Db 301 TTGTCTTGAAGAACTCAAGAGATGACATGACAGAGATCTTCCGGCTTTGCAAACTGC 360

QY 2117 TGGAAACCTGAGTGGAGGGGTGTCTTCTCTAGATTCAAGACCAAAAGATTAATAAT 2176
Db 361 TGAAGCATATGCTATGAGAGTGTCTTCTCTAGATTCAAGACCAAAAGATTAATAAT 420

QY 2177 TTTAACTCGGGAATCAATATGATGATCTATGTTTGAAGTTACTTGAAGAGTGGAGG 2236
Db 421 TCAAGCTTGAAGATTAATATGATGATCTATGTTTGAAGTTACTTGAAGAGTGGAGG 480

QY 2237 TAGTTCAAGGGTCTCTTATAGTGTCTGCTCAACTATGCGAAGATGAGCCGAG---- 2292
Db 481 GACAGAGTGTCTTCACTTGTGCTGACCCAGCTAATAGAGAGATGAGTGGAGAGCTTA 540

QY 2293 -----CATGTAAAGCTAGTGTATGAGGCACTGCAAGATTTTCTTCTT 2338
Db 541 CTGCAAGTTCTTGAATCAATGAAAGCTATGCAATTCAGGCAATGCAAGAGTGTCTTC 600

QY 2339 CCGCTAATACAGA 2351
Db 601 TTCGTGCTCAGA 613

RESULT 14
CR486014/c
LOCUS
DEFINITION
Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago
ACCESSION
CR486014 GI:48647590
VERSION

KEYWORDS
SOURCE
ORGANISM

GSS.
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.

REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (10-JUN-2004) Genoscope - Centre National de Sequencage ;
BP 101 91006 EVRY cedex - FRANCE (E-mail : seqrel@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

FEATURES
source
Location/Qualifiers

1. .751
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="Jemalong A17"
/db_xref="taxon:3880"
/clone_lib="MTB2"
/note="Vector: pBelOAC11 ; Site_1: HindIII ; Site_2: HindIII ; Cook, D.R. and Kim, D.J.-Genoscope sequence ID : mth2-155019FW2"

ORIGIN

Query Match 7.2%; Score 265.4; DB 9; Length 751;

Best Local Similarity 64.1%; Pred. No. 5.4e-57;

Matches 476; Conservative 0; Mismatches 216; Indels 51; Gaps 3;

QY 1491 ATGATTAACGTGTGGAAGAAACCTAAATGTTTAAGCGGTGTGGAGATTTTGTGCTG 1550
Db 744 ATTAACAATCGAGGCGGAGAGAAAGTCTCCAAAGTGTCCCAATTCCTGTGGCAG 665

QY 1551 TTGAGAGAGTGAAGCATCAGCTTGTGTGGGGTGTGACCCGTGAAAGTTATGAAATG 1610
Db 684 TTGAGAGTGAAGAGACAGCAGCAATTCGTGGAGTTTCAAGCTGAGATTTATGAAATG 625

QY 1611 AGGCGTTTTCGAATGACAGCTGTGAGAGGTAACAGTTTATGATA-----CCTTTT 1664
Db 624 AGCATTTCCTGCAATGAAAGTGTCCGAACAGGTTCATAGAAATATGATGATGTT 565

QY 1665 TTTAATTTCTTATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1713
Db 564 TTTAATTTATTTCTTCCGCCCCCTTATGTTTGTATGAAAGTTTTACATTAATTC 505

QY 1714 -----GTGTGTGTGAGTGTATCTTTTGTAGCTAC 1745
Db 504 CTTTCATTAATTTCTTTTATCTCCAAATCTTTTTTTAAAGTGAACCTTTTGTAGCAAC 445

QY 1746 CCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTTGTGGC 1805
Db 444 ACCTACCAATATTCAGAGTGAAGTTTGAAGCTTGAAGCTTGTGCACTTGTGTGGC 385

QY 1806 TCAAGCTTTATTTGTAAGAGCCACACCTTTTACAGATGCTGATPAGCAATTCAGCA 1865
Db 384 ACAAGCTTTGTAAGTAAAGCCACATCTTAACCAAGATGCTGATTAATTAATTCATCA 325

QY 1866 ACTTCAGAGGCTTAAGTATAGGCTATGAGAGATTCCTGGAAGTGTATGATAC----- 1919
Db 324 ACTTCAACAACTTAAGTATCAATATGAGAGATGCTCCTCTGTATTAATCTCCATGGA 265

QY 1920 ACCGAATTAATTTGAGATAGACTTCGCTTGAAGAGGAGCTTGTGCACTGCTTAAGG 1979
Db 264 GATGAGAGAGAGAGAGAGTTGATTTGCAATTAAGAAAGGGGTGTGTGCACTGCTTGTGG 205

QY 1980 CAAAGTTGATGAATGCCGATGTGTGGCTTGAAGCAGTGAAGATTCACATATATGAA 2039
Db 204 GAGCTTGAATCAATGTGATCAATGTGTGGCTCGATGAGCAGCTCACTTAATGAA 145

QY 2040 TCCAGCTATTTGGAATTTGTTTGAAGATTCGAATCGATGAGCAATGATGATCTCC 2099
Db 144 CCCATCATTAATGACCTTAATTAAGAAACGAAAGGATGAGCAAGATGATCTTC 85

QY 2100 TGACATATGCAATTTGTTGAAACCTGGTGGACAGGGGTTGCTTTCTTACAGTTCAAGAGA 2159
 DB 84 TGACCTCTGTAATATGTTTGGAGACATGTTGATGAGGTGGGTTTCCCTAGATTGAAGA 25
 QY 2160 CACCAAGATTAATAATTTAAAC 2182
 DB 24 TACTAAGAGCAAACTTTAAAC 2
 RESULT 15
 COL17047 849 bp mRNA linear EST 16-JUN-2004
 LOCUS GR_EB019K01.r GR_Eb Gossypium raimondii cDNA clone GR_EB019K01
 DEFINITION 3', mRNA sequence.
 COL17047
 ACCESSION COL17047.1 GI:48815734
 VERSION COL17047.1 GI:48815734
 KEYWORDS EST.
 SOURCE Gossypium raimondii
 ORGANISM Gossypium raimondii
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 849)
 K1m,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
 Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
 Wing,R.A.
 Global assembly of Cotton ESTs
 TITLE Unpublished (2004)
 JOURNAL Contact: Rod A. Wing
 COMMENT Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: http://genome.arizona.edu
 Plate: 019 row: K column: 01.
 Location/Qualifiers
 1..849
 /organism="Gossypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_EB019K01"
 /tissue_type="floral"
 /dev_stage="3 to +3 DPA"
 /lab_host="MD10B"
 /clone_lib="GR_Eb"
 /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
 EcoRV; library made by Invitrogen with RNA supplied by
 Wendle lab. Directional cloned into NotI-EV. Clones
 plated/picked by AGI. More glycerol clones held in -80."
 ORIGIN
 Query Match 6.9%; Score 254; DB 7; Length 849;
 Best Local Similarity 58.7%; Pred. No. 5e-54;
 Matches 518; Conservative 0; Mismatches 325; Indels .39; Gaps 3;
 QY 1828 CCACACCTTTTACAGATGCTGATGAAGCAATTCAGCACTTCAGCGGCTAAGTAATG 1887
 DB 1 CCTCATCTCATTAAGATGCTGATTAACCTCTTCCAGAGCTTCAGACACTAAGTAACA 60
 QY 1888 GCTATGAGAGATTCCTGGAGATGTTGATGATACAGGAAATATGAGAGATAGACTCGGT 1947
 DB 61 ACTTTTGAAGACAGTGTCTCTCTATGCTCTGTGCGAAGCCGTGAGATGACTTTGCT 120
 QY 1948 CTAGAAAGGGACTCTGTGCACTGCTTAAGCAAAAGTTGATGATGCCGTATGTGTTG 2007
 DB 121 TTGAGAGGGGCTCTGTTCATGCTGTGGGGAGCTTGATGATGCGCTTCGAGTTG 180
 QY 2008 GGCTTAAGCACTGAGATTCACATATGGAATCAGCTATTTGAGATTGTTTGGAG 2067
 DB 181 GGCTTAAGCACTGATGCTCCCTTAAGAAATACATCTATTTGAGAAATTTGCTTGGAA 240
 QY 2068 AATCAATCGTATGACATATGATCTCCCTGGAATGATGCAAAATTTGTTGAAACCTGG 2127

DB 241 AACTCAAGAGATGACGATGACAGAGATCTCCCGGGCTTTGCAAACTGCTGAGGCAATGG 300
 QY 2128 TTGGCAGGGGTGTCCTTCCAGGTTCAAGAGACACCAAAAGATTAATAATTTAACTCGGG 2187
 DB 301 CTATGAGAGGTGTTTCTTCTAGATTAGACACCAAAAGATTAATCAATTTCAACTTGA 360
 QY 2188 GACTACTATGATGATGCTTATGATGATTTGAGATTACTTGAAGAGTGAAGTTCAGGGT 2247
 DB 361 GATTATTATGATGATGCTTATGATGATTTGAGATTATTTGAAAGGCTTGAAGGAGCAGTGGT 420
 QY 2248 TCTCCTTTAGTGTGCTGCTGCACTATGAGCAAGATTGAGCCGACATGTAAGTAACTAGT 2307
 DB 421 TCACCTTTGCTGAGCGGACGCTATATGATGAGATGAGGCTGACGAGGCTTCACTGCACTT 480
 QY 2308 GCTATGAGGACATGACAGAAAGTTTCTCCCGGCTATACAGATAGAAACTCGGCTGAA 2367
 DB 481 GAT-----CATGTAAAGGCTAGTGCATTCAG 507
 QY 2368 CCCAAGATGTCAGAGACAGTGTATGATGATCTGTTGTTAACTATGAGCCGT 2427
 DB 508 GCATTCAGAAAGTGTTCCTCTTCGTGCTCAGAGAGACTGCGCAGATCAATTAATGAT 567
 QY 2428 GATGTCAGCCTGCTGCTTTATTTGCAAGAGCTGTAAAGCCCTGTGAAACTTTGAAACT 2487
 DB 568 GTGAGATGAACAAACTTCTCTCTGTGAAAGTGAAGAGACTCTTGAAGAACTGATCAA 627
 QY 2488 AATGATATGCAATTCGAGCTGGGCTCTC---AGAGGTAGCGTTGATGAACTACTGTT 2544
 DB 628 GAGATTCGTCATCTTCTAGCTGAGGTTCTGGAATTAAGCAGTTTGAAGGAAATGATGAA 687
 QY 2545 GAAATGTCCTGCTGATGATGTTAAAGAGCAAGTGTGAAGATCTTATGCTGTGTGTG 2604
 DB 688 GAGAAACATTAATGACAAATTAAGATGCCAGTGTCAAGATCATGTGCTGAGTGTG 747
 QY 2605 GCAATTCAGCTGATTTCACTGTTCAAGCCAGAGTATTT---TCTTAAAGCAGCTCATCT 2661
 DB 748 GTAATGAGTGAATGACATTTGCTGTGAAGTTTATCTGTTAAATTTAGTTCAATCA 807
 QY 2662 TTTCAGCAAGATATGCTTTCTTATGAGAACTGATGTC 2703
 DB 808 GTGACCGGTAAAGGCAATCAGTCAGCTATGCAACTGATGTC 849

Search completed: June 10, 2005, 00:53:15
 Job time : 11203 secs

THIS PAGE BLANK (USPTO)